PA

It's All in the Name: A Character-Based Approach to Infer Religion

Rochana Chaturvedi[1] and Sugat Chaturvedi[2]

[1] Department of Computer Science, University of Illinois Chicago, Chicago, IL, USA. Email: rchatu2@uic.edu
[2] Science Policy Research Unit (SPRU), University of Sussex, Sussex House, Falmer, Brighton BN1 9RH, UK. Email: sc2057@sussex.ac.uk

Abstract

Large-scale microdata on group identity are critical for studies on identity politics and violence but remain largely unavailable for developing countries. We use personal names to infer religion in South Asia—where religion is a salient social division, and yet, disaggregated data on it are scarce. Existing work predicts religion using a dictionary-based method and, therefore, cannot classify unseen names. We provide character-based machine-learning models that can classify unseen names too with high accuracy. Our models are also much faster and, hence, scalable to large datasets. We explain the classification decisions of one of our models using the layer-wise relevance propagation technique. The character patterns learned by the classifier are rooted in the linguistic origins of names. We apply these to infer the religion of electoral candidates using historical data on Indian elections and observe a trend of declining Muslim representation. Our approach can be used to detect identity groups across the world for whom the underlying names might have different linguistic roots.

Key words: religion inference, person names, South Asia, machine learning

1 Introduction

Names are associated with group identity around the world. An important marker of identity is religion which shapes preferences, attitudes, and political and economic outcomes (Guiso, Sapienza, and Zingales 2003; Iyer 2016). In India too, it is associated with socioeconomic status, health outcomes, electoral behavior, and conflict (Bhalotra, Valente, and Van Soest 2010; Chhibber and Shastri 2014; Iyer 2018). Despite its salience, there is a lack of fine-grained data on religion in South Asia.[1] Therefore, much of the research relies partly or wholly on manual classification. For example, Sachar *et al.* (2006) infer religion using person names to highlight economic and social deprivations faced by Indian Muslims. Others use electoral candidate names to examine the effect of co-religiosity on voting behavior and of Muslim representation on education and health outcomes of constituents (Bhalotra *et al.* 2014; Heath, Verniers, and Kumar 2015). Similarly, Field *et al.* (2008) use names in electoral rolls to examine the effect of residential segregation on Hindu–Muslim violence during the 2002 Gujarat riots. This dependence on manual classification limits studies on religious demography to coarse or small scale analyses.

In this paper, we bridge this gap by training character-sequence-based machine-learning models that infer religion using person names alone. While our methods are more generally applicable for predicting other markers of group identity such as race, gender, ethnicity, caste, and nationality, we demonstrate their strength by inferring religion in the Indian context as a case in point.[2] In India, names are well known to signify religious identity. This is evident in Gaikwad

1 While the Indian Census collects this at the individual level, it is publicly released as an aggregate only upto the sub-district level. In contrast, in developed countries such as the United States, microdata on race is readily available—facilitating studies on racial discrimination and residential segregation (Ananat and Washington 2009; Cutler, Glaeser, and Vigdor 1999).

2 We show in Section A of the Supplementary Material that these methods can also be used to infer race/ethnicity in the United States and substantially improve Black voter share estimates over the existing approaches.

Political Analysis (2024)
vol. 32: 34–49
DOI: 10.1017/pan.2023.6

Published
23 March 2023

Corresponding author
Sugat Chaturvedi

Edited by
Jeff Gill

and Nellis (2017) who assign Hindu or Muslim sounding names to fictitious internal migrants and elicit attitudes of natives toward them in a face to face survey in Mumbai. Moreover, name lists along with precise addresses or locations are often publicly available. Sources such as electoral rolls, below poverty line (BPL) lists, land records, and beneficiary lists of social security programs such as job cards for Mahatma Gandhi National Rural Employment Guarantee Act (MGNREGA), Swachh Bharat Mission (SBM), etc. provide multiple related names and locations for millions of households but do not disclose religion. Our work, therefore, can be used to construct individual level datasets incorporating names to study religious demography and uncover discrimination in the allocation of targeted welfare programs.[3] Inferring religion from names on social media platforms that lack demographic attributes can also guide network and sentiment analysis, and detect religious polarization.

Currently the only viable alternative to manual classification of names into religion is provided by Susewind (2015), who uses a string matching algorithm to predict religion based on a reference list. However, being a dictionary based method, the algorithm suffers from low coverage; it cannot classify unseen names and is not resilient to spelling variations. We show that character-based machine-learning models outperform the existing work while being orders of magnitude faster. Our models can also classify unseen names with high accuracy and account for spelling variations.

Due to their distinct linguistic origins, Muslim and non-Muslim names are particularly interesting. While Classical Arabic is the liturgical language of Islam, Sanskrit is the principal liturgical language of Hinduism. Buddhism, Sikhism, and Jainism are also rooted in Sanskrit or Indic languages Pali, Punjabi, and Magadhi Prakrit, respectively. The distinct orthographies of the linguistic roots manifest in person names too. Islamic names are derived from Classical Arabic, Persian, and Turkish whereas non-Muslim names are rooted in Sanskrit or Dravidian languages (Emeneau 1978; Schimmel 1997). We explain the classification decisions of one of our models and systematically uncover prominent linguistic differences between Muslim and non-Muslim names by applying the layer-wise relevance propagation (LRP) technique from the field of explainable artificial intelligence (Bach *et al.* 2015). The model associates typical character patterns, meaningful prefixes, and suffixes in Classical Arabic and Sanskrit/Dravidian languages with Muslim and non-Muslim names, respectively.

The politics and economics of Hindu–Muslim relationship is also of interest to social scientists (Bhalotra *et al.* 2021; Mitra and Ray 2014; Nellis *et al.* 2016). Indian Muslims, despite being a sizable minority (14% of population), are persistently under-represented in office at both the state and national levels. Consequently, they suffer economic and social backwardness (Sachar *et al.* 2006). Compared to the Hindu majority, Indian Muslims have lower access to publicly provided goods such as tap water (64% vs. 70% for Hindus; NSS 69th round, 2012) and healthcare (2.1% vs. 5.1% for Hindus; NFHS-3), and have lower education attainment (Kundu 2014). We apply our model to candidates' names from large-scale data on national- and state-level elections in India and highlight the puzzling trend of declining Muslim representation despite a consistent increase in their population share.

2 Names and Group Identity

A substantial literature leverages names as signals of group identity. Several studies use fictitious resumes to find evidence of labor market penalty associated with African-American or foreign sounding names in North America (Bertrand and Mullainathan 2004; Oreopoulos 2011) and Arabic names in France (Adida, Laitin, and Valfort 2010). In response, the discriminated minorities might

3 Chaturvedi, Das, and Mahajan (2021) use one of our models to infer the religion of over 25 million households in rural Uttar Pradesh, India. They examine how the effect of gender quotas on public good provision varies with Muslim share in village councils.

change their names to signal an intent to assimilate (Algan *et al.* 2022; Biavaschi, Giulietti, and Siddique 2017; Fouka 2019) or assert their cultural identity via names (Fouka 2020).

Our work is related to the literature that infers group identity from names. Harris (2015) uses geocoded person names to estimate local ethnic compositions in Kenya by modeling the ethnic proportions of each unique observation in a surname list. Elliott *et al.* (2009) introduce Bayesian Improved Surname Geocoding (BISG) which infers an individual's race given their surname and geolocation using Bayes' rule.[4] Imai and Khanna (2016) improve upon this by combining surnames and geolocation with age, gender, and party registration using Florida voter registration data.[5] BISG requires information on racial/ethnic compositions at each precise geolocation. This might be useful when identity groups are spatially segregated.[6] Given residential segregation along religious lines in India, geographic information might be useful in this context as well.[7] However, geocoding can be expensive and is often inaccurate. For reference, Google API costs $5.00/1,000 addresses. Clark, Curiel, and Steelman (2021) use ESRI 2013 street address geocoder which was unable to geocode 4.6% of addresses in Georgia, United States. Moreover, outside the United States—especially in developing countries—information on group composition may not even be available at geographically precise levels making geocoding infeasible. In such contexts, inferring group identity only from names may be useful. When geocoding is feasible, our character-based models can be incorporated within the existing BISG packages to improve performance over the surname dictionaries.

In contrast to race/ethnicity inference, religion inference has only received a limited attention. The case of religion is distinctive, especially given that people can have multiple races or ethnicities when they descend from more than one racial group, but only a single religion even in case of interfaith marriages.[8] Moreover, in India, interfaith marriages are rare. A second distinction that makes religion inference from names interesting is that, unlike race, religion entails codified sets of beliefs and practices for its adherents. This is also reflected in personal names for which there are prescriptive guidelines across different religions. For example, in South Asia, Islamic naming guidelines prescribe Arabic names taken from the Quran and recommend avoiding resemblance to Hindu names (Metcalf 2009). On the other hand, there has been a shift in naming conventions across racial groups in the United States over the past few decades. Fryer Jr and Levitt (2004) discuss how Blacks living in predominantly White neighborhoods increasingly adopt White sounding names than Blacks in racially segregated neighborhoods. Therefore, we expect names across religions to remain more distinctive than across racial/ethnic groups.

Our approach is also related to several papers in machine learning that infer demographic attributes from names. Early papers in this literature almost exclusively use dictionary-based methods which suffer from lack of coverage on unseen names or spelling variations (Mateos 2007). To address this, others use sub-name features such as character *n*-grams, prefixes, suffixes, and phonetic patterns. These are used to infer nationality, gender, and ethnicity using hierarchical decision trees and hidden Markov models (Ambekar *et al.* 2009), Bayesian inference (Chang *et al.* 2010), multinomial logistic regression (LR) (Torvik and Agarwal 2016; Treeratpituk and Giles 2012), and support vector machine (SVM) (Knowles, Carroll, and Dredze 2016). Lee *et al.* (2017) and

4 This requires the assumption that an individual's surname is orthogonal to their location given their race.
5 Müller-Crepon and Hunziker (2018) do not use names, but use spatial interpolation and machine learning to obtain ethnic settlement patterns at the local level.
6 In our experiments on race/ethnicity inference in the United States, incorporating geographic information increases the macro-average F_1 scores of our character-based models by 2.4–4.5 percentage points (results available on request).
7 This might especially improve identification of Sikhs, Buddhists, and Christians who are highly segregated across Indian sub-districts. The dissimilarity index—a commonly used measure of segregation interpreted as the proportion of a group that would have to relocate to have even spatial distribution of that group—across sub-districts for Hindus, Muslims, Christians, Sikhs, Buddhists, and Jains is 42.65%, 44.15%, 72.50%, 88.66%, 83.09%, and 65.31%, respectively (authors' calculation based on the 2011 Census).
8 For example, according to the 2020 U.S. Census, 10.2% people report belonging to multiple racial groups.

Wood-Doughty *et al.* (2018) avoid manually crafting sub-name features and use neural networks that learn these features automatically from the character sequence in names. While Bayesian inference and hidden Markov models are generative classifiers that model the probability of an output class, decisions trees, LR, SVM, and neural networks are discriminative classifiers that learn a boundary separating the classes and are well suited to classification tasks. Our experiments reaffirm that discriminative classifiers outperform dictionary-based and generative model baselines.

3 Data

3.1 REDS

We use the Rural Economic & Demographic Survey (REDS) data collected by the National Council of Applied Economic Research to train our models. It constitutes a nationally representative sample of over 115,000 rural households from 17 major Indian states surveyed in 2006. We use the respondent's and their parent/spouse's name and self-reported religion. We label a person as Muslim or non-Muslim and split the data into training, validation, and test sets in the ratio 80:10:10.[9]

3.2 U.P. Rural Households

One concern with self-reporting in REDS could be that some people might not accurately reveal their religion, for example, due to fear of persecution. This might be a source of noise, and we expect that our models would have been even more accurate if there was no misreporting. Therefore, we use a second test set to further validate our models. Due to a lack of publicly available datasets mapping names to religion, we annotate the religion of 20,000 randomly selected household heads from a dataset comprising over 25 million households in rural Uttar Pradesh (U.P.)—the largest state of India.[10] Hindus (comprising 83.66%) and Muslims (15.55%) are the predominant religious groups in rural U.P. and form over 99.2% of the population. Therefore, the annotators classify the religion as either non-Muslim (largely comprising Hindus) or Muslim.

The annotations are done independently by the two annotators using the names of household heads and their parent/spouse. The inter-annotator agreement rate is 99.91% (Kohen's Kappa κ = 0.9959) indicating that names strongly reflect perceived religion.[11] We further validate the veracity of annotations by manually classifying randomly selected 1,000 person names in REDS as Muslim or non-Muslim. The annotations are accurate for 99.2% cases indicating large overlap between self-reported religion and annotations.

Table 1 shows descriptive statistics for both the datasets. REDS contains nearly 99,000 unique names ($\approx$ 86% of observations) while the corresponding figure is over 12,000 ($\approx$ 62%) for U.P. Rural Households dataset. The average name length is 15.6 and 8.8 characters in the two datasets, respectively. The shorter name length for the U.P. Rural Households dataset is due to nearly 60% observations containing information only on an individual's first name. In contrast, REDS includes both first and last names for 95% individuals.[12] The religious composition in the REDS data closely mirrors the national level rural composition.[13]

Figure 1 shows relative character frequency distributions representing the ratio of average frequency of each character to average name length across a religious group in REDS. The alphabets "F", "Q", and "Z" are characteristic of Muslim names. They represent phonemes [f], [q], and [z],

9 Replication data and code for this study are available in Chaturvedi and Chaturvedi (2023).
10 These names are in public domain and scraped from https://sbm.gov.in.
11 The disagreements were resolved in consultation with Sanskrit and Arabic experts.
12 This also holds when we restrict the REDS sample to rural U.P. and for both Muslims and non-Muslims. Thus, the discrepancy in name lengths is primarily explained by survey design rather than regional difference.
13 According to the 2011 Census, the religious composition in rural India is as follows: Hindus: 82.05%; Muslims: 12.41%, Christians: 2.00%; Sikhs: 1.79%; Buddhists: 0.58%; and Jains: 0.11%. Overall, in India, Hindus are the majority comprising 79.8% of the total population. Muslims (14.23%) are the largest minority followed by Christians (2.30%), Sikhs (1.72%), Buddhists (0.70%), and Jains (0.37%).

Table 1. Descriptive statistics.

	REDS	U.P. rural
# Unique names	98,853	12,342
# Unique characters	27	27
Average name length	15.57	8.82
Longest name length	40	29
% Buddhist	0.42	–
% Christian	2.46	–
% Hindu	84.47	–
% Jain	0.30	–
% Sikh	3.22	–
% Muslim	9.13	13.32
% Non-Muslim	90.87	86.69
No. of obs.	115,180	20,000

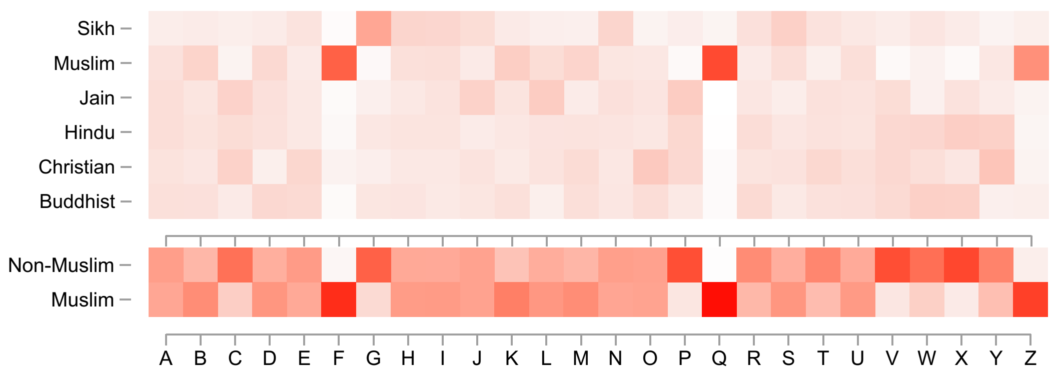


Figure 1. Relative character frequency heatmaps for REDS data.

respectively, that do not exist in the Sanskrit phonemic inventory. On the other hand, "P", "V", and "X" are rare among Muslim names owing to the absence of phonemes [p], [ʋ], and [ʂ] in Classical Arabic. Hindu, Sikh, Jain, and Buddhist names have similar distributions owing to their common linguistic origins.[14]

4 Models

We make predictions using single and two names (i.e., primary and parent/spouse's name) in each household. We preprocess the raw data by upper-casing and removing special characters, numbers, and extra spaces. For single name models, we also include parent/spouse's name as a primary name to enrich our training set as it is highly likely that they share the same religion. Since REDS is a nationally representative survey, we keep duplicates to account for frequency of each name within a religion. We describe our models below and defer technical details to Section B of the Supplementary Material.

4.1 Baseline: Name2community

We use a dictionary-based classification algorithm Name2community proposed by Susewind (2015) as our first baseline. The algorithm first counts the frequency of each name part (i.e., the

14 People who convert to Christianity in India often retain their original names. According to the 2019–2020 PEW Research Center survey comprising 29,999 adults, 0.4% of the respondents converted to Christianity (Sahgal *et al.* 2021). Overall 98% respondents continue to follow the religion in which they were raised.

first name, last name, etc.) within a reference list specific to each religion based on spelling and pronunciation. These two frequencies are combined to obtain a certainty index for each name part for each religion. These indices are then aggregated over all the name parts to get the certainty index for the entire name. Finally, each name is assigned the religion having the highest certainty index.[15]

4.2 Baseline: Language Models

For our second baseline, we follow Jensen *et al.* (2021) who train language models to infer *religiosity* from Indonesian names. A language model computes the probabilities of *n*-grams from a training corpus. It then uses a method such as maximum likelihood estimation to predict the probability of the next character or word in a given sequence. Language models have previously been used for a distinct task of language and dialect identification from text pioneered by Cavnar and Trenkle (1994) and improved upon by Vatanen, Väyrynen, and Virpioja (2010) and Jauhiainen, Lindén, and Jauhiainen (2017, 2019a).[16] We train two separate language models LM_M and LM_{NM} on the set of Muslim and non-Muslim names, respectively. We then compute perplexity—a standard metric for evaluating language models—of both the models for a given name. Perplexity measures how surprised a model is on seeing a name in the test set. Therefore, we classify a name as Muslim if the perplexity score of LM_M is less than that of LM_{NM} for that name and non-Muslim otherwise.[17]

4.3 Bag-of-*n*-Grams Models

For bag-of-*n*-grams models, we first convert each name to its character *n*-gram feature representation using term frequency-inverse document frequency (TF-IDF).[18] TF-IDF captures the importance of each character *n*-gram (or token) in a document normalized by its importance in the entire corpus without taking into account its relative position in the document. We then use linear SVM and LR classifiers with L2 regularization to predict religion from these feature vectors. Since the classes are highly imbalanced, we use balanced class weights.

4.4 Convolutional Neural Network

We experiment with several neural network architectures popular in text classification based on Convolutional Neural Network (CNN) and Long Short-Term Memory network (LSTM). We find that character-based CNN gives better performance on our task. Originally designed for computer vision, CNN is known for its ability to extract important local features using far fewer parameters compared to other neural models (LeCun *et al.* 1989). Its architecture is also highly parallelizable making it faster. CNN has attained much success in natural language processing research since Collobert *et al.* (2011) who apply it to multiple NLP tasks for improved speed and accuracy.[19] Our model takes character sequence in a name as input using an architecture similar to Zhang, Zhao, and LeCun (2015) and outputs the probabilities of the name belonging to each religion.[20]

5 Model Decisions and Linguistic Roots

Machine-learning models are often black boxes—they are good at predicting an outcome, but the reason for their predictions is unknown. It is important to explain how a model arrives at a decision

15 The implementation is available at https://github.com/raphael-susewind/name2community.
16 See Jauhiainen *et al.* (2019b) for a detailed survey.
17 We also experiment with Naïve Bayes classifier with smoothing. It does not perform well in our case. The results are available on request.
18 This improved performance over handcrafted sub-name features as described in Knowles *et al.* (2016).
19 Also see Torres and Cantú (2022) who introduce CNN to social scientists and show an application by digitizing handwritten information from vote tallies.
20 For choosing hyperparameters, we perform Bayesian search followed by manual tuning. Tables 8 and 9 in Section C of the Supplementary Material describe the search space and the selected hyperparameters.

Table 2. Heatmaps for Muslim and non-Muslim names using LRP on distinctive REDS test set names.

s.no.	True Muslims	True non-Muslims
1.	{ A B D U L } { M O H A M M E D } { H A N I F }	{ A R E T I } { P A L E S W A R A } { R A O }
2.	{ A K B E R } { A L I } { S I D I Q I }	{ B A L D E V } { B H A I } { T H A K O R }
3.	{ A L L U D D I N } { M I Y A }	{ D E V I R A M } { S A H U }
4.	{ F A R U Q } { H U S S A I N } { P I N J A R A }	{ D H A R A M } { P A L } { S I N G H } { J A T }
5.	{ J A F E R U L L A H } { K H A N }	{ K A N N I A P P A N } { G O V I N D A S A M Y }
6.	{ S K } { H A F I J U L } { R A H A M A N }	{ K U M A R A S A M Y } { T H N U } { P I L A I }
7.	{ M D } { I Z B U L }	{ R A M } { P A L } { J A T A V }
8.	{ H A S I N A } { B E G A M } { K H A N }	{ G U R D E V } { K A U R }
9.	{ J U B A I D A } { K H A T U N }	{ K E T A } { V E N K A T A L A X M I }
10.	{ M A I M U N } { B I B I }	{ P R A T I B H A } { S I N G H } { S I R O H I }

to assess its validity and generalizability, and to foster trust in it. We apply LRP on the REDS test set to identify the character patterns distinguishing Muslim and non-Muslim names in India.

LRP maps the output of a classifier back to the input characters and computes the contribution of each input character to the final prediction of a machine-learning model. In our context, it answers what character patterns make a name Islamic or non-Islamic. Arras *et al.* (2017) apply LRP to text data and show that though both SVM and CNN models performed comparably in terms of classification accuracy, the explanations from CNN were more human interpretable. Therefore, we study the decisions of our CNN model using the LRP implementation of Ancona *et al.* (2018).[21]

Table 2 reports LRP heatmaps showing classification decisions of CNN model on distinctive names from REDS test set. Characters with positive relevance scores with respect to Muslim class are labeled red, while those with negative relevance scores are blue, that is, they have positive relevance for non-Muslims. The left panel shows examples of correctly classified Muslim names. The LRP relevance scores are able to identify phonemes characteristic in Classical Arabic such as "F" (column 1, examples 1, 4–6), "Q" (column 1, examples 2 and 4), "Z" (column 1, example 7), and "KH" (column 1, examples 8 and 9). Meaningful suffixes such as "UDDIN" (column 1, example 3) meaning "(of) the religion/faith/creed" that are highly characteristic of Arabic names are also detected as relevant for the Muslim class. The right panel shows correctly classified non-Muslim names. The characters "P" (column 2, example 1), "V" (column 2, example 2), and "X" (column 2, example 9) are highly relevant for non-Muslims.

The relevance of a character toward a class also depends on its context. For example, the neutral character "D" is highly relevant to the Muslim class when it is a part of "UDDIN", while it becomes highly relevant to non-Muslim class when it forms the word "DEV" (meaning god in Sanskrit) (column 2, examples 2, 3, and 8).[22] We notice that the relevance of the characters "{" and "}" signifying the beginning and end of a name part, respectively, is also modulated by the character sequences following and preceding them.

In Table 3, we show 10 most relevant unigrams, bigrams, and trigrams for both the classes conditional on n-grams not being rare, that is, occurring at least 25 times in the test set. For this, we apply LRP on all test set names and average the relevance scores of each n-gram. The linguistic differences discussed in Section 3 are indeed systematically captured by the model. Unigrams "F", "Q", and "Z" are most predictive of the Muslim class, whereas "X", "V", and "P" are most relevant

21 https://github.com/marcoancona/DeepExplain.
22 We see similar contrast for (1) "L" in "UL" (meaning "of" in Arabic) vs. "PAL" and (2) "M" in "MOHAMMED" vs. "RAM" (the name of a Hindu deity).

Table 3. Most relevant *n*-grams among Muslim and non-Muslim names.

Unigram		Bigram		Trigram	
Muslim	Non-Muslim	Muslim	Non-Muslim	Muslim	Non-Muslim
F	X	F}	PR	F}{	PRA
Q	V	IF	IV	{SK	DEV
Z	P	AF	GW	SAB	{PR
B	W	FI	EV	SK}	GHO
H	G	FA	VV	BEG	SIV
J	C	B}	SW	DDI	EV}
U	T	KH	MP	AB}	EGH
–	Y	FU	EP	KH}	DEY
–	R	DD	LD	FAR	PAL
–	O	ZA	V}	BI}	VVA

for the non-Muslim class. The bigram "KH" corresponding to phoneme [x] is highly relevant for the Muslim class. The character positions are also important in a name part. We find that "F}" and "B}" are highly relevant to the Muslim class implying that the characters "B" or "F" at the end of a name part characterize Muslim names. On the other hand, the bigram "PR" is a distinguishing feature of the non-Muslim class, especially at the beginning of a name part, denoted by the trigram "{PR". This is meaningful as "PR" is a Sanskrit prefix which when added to an adjective or a noun accentuates its quality. Similarly, the bigram "VV" has positive relevance for non-Muslim class as it forms part of the Dravidian honorific suffix "AVVA" added to female names. The trigram "DDI" is considered highly relevant by our model and forms part of the suffix "UDDIN" in Arabic names. These examples illustrate that LRP relevances are very reliable at finding meaningful character *n*-grams that distinguish the two classes and highlight the linguistic differences depicted by the names.[23]

6 Results

We report the results in Table 4 for both REDS and U.P. Rural Households test sets. For evaluation, we use Precision (P), Recall (R), and their harmonic mean (F_1) defined as follows:

$$P = \frac{TP}{TP+FP}, \quad R = \frac{TP}{TP+FN}, \quad F_1 = \frac{2*P*R}{P+R},$$

where TP = #True Positives, FP = #False Positives, and FN = #False Negatives. In our context, precision measures what percentage of individuals predicted to have a particular religion actually belong to that religion. Recall measures what fraction of members actually belonging to a religion is classified under that religion. The two metrics are especially useful for imbalanced classes. Precision will be low if there are too many false positives and recall will be low in case of too many false negatives. Their harmonic mean F_1 captures the trade-off between the two types of errors.

Panel A of Table 4 shows the results when predicting religion using only a single name. Name2community can only classify less than two-thirds of observations in both the test sets. The scores in the table are based on the observations classified unambiguously. We also experiment by assigning the majority religion to the ambiguous predictions of Name2community for

23 See Section D of the Supplementary Material for a discussion of which name part is most useful in identifying the religion based on relevance scores.

https://doi.org/10.1017/pan.2023.6 Published online by Cambridge University Press https://doi.org/10.1017/pan.2023.6 Published online by Cambridge University Press

Table 4. Results on test sets. Standard errors are reported in parentheses. The highest score for a metric within a panel is marked in bold. The evaluation for Name2community and Language Model is based on names classified unambiguously.

		REDS						U.P. Rural Households					
	Models	**Coverage**	F_1	**Muslim**		**Non-Muslim**		**Coverage**	F_1	**Muslim**		**Non-Muslim**	
				P	**R**	**P**	**R**			**P**	**R**	**P**	**R**
Panel A: Single name	Name2community	65.28	93.39	90.08	85.82	98.65	99.10	57.34	**93.11**	**92.58**	83.20	97.98	**99.19**
				(0.56)	(0.53)	(0.17)	(0.17)			(0.47)	(0.42)	(0.15)	(0.15)
	Language model	97.99	89.78	76.87	86.75	98.67	97.42	97.89	85.27	66.74	85.40	97.71	93.61
				(0.51)	(0.57)	(0.17)	(0.18)			(0.41)	(0.52)	(0.18)	(0.20)
	Logistic regression	100	95.23	89.28	93.53	99.35	98.88	100	90.24	79.70	87.01	97.98	96.60
				(0.37)	(0.39)	(0.12)	(0.12)			(0.37)	(0.40)	(0.15)	(0.16)
	SVM	100	95.64	90.15	**94.10**	**99.41**	98.97	100	91.45	82.70	**87.95**	**98.13**	97.17
				(0.35)	(0.37)	(0.11)	(0.12)			(0.36)	(0.38)	(0.14)	(0.15)
	CNN	100	**95.86**	**94.62**	90.39	99.04	**99.49**	100	90.67	90.19	77.99	96.69	98.70
				(0.36)	(0.35)	(0.11)	(0.11)			(0.41)	(0.36)	(0.15)	(0.14)
Panel B: Concatenated	Name2community	72.55	94.24	91.69	87.38	98.77	99.23	72.72	93.76	92.51	85.67	98.07	99.05
				(0.50)	(0.47)	(0.15)	(0.15)			(0.39)	(0.36)	(0.14)	(0.13)
	Language model	97.82	92.67	85.63	87.75	98.79	98.55	96.19	91.21	84.39	84.94	97.79	97.70
				(0.46)	(0.47)	(0.15)	(0.15)			(0.38)	(0.38)	(0.15)	(0.15)
	Logistic regression	100	97.32	94.63	95.62	99.56	99.46	100	96.49	94.93	92.90	98.91	99.24
				(0.29)	(0.29)	(0.09)	(0.09)			(0.24)	(0.24)	(0.09)	(0.09)
	SVM	100	**97.33**	94.13	**96.19**	**99.62**	99.40	100	**96.56**	94.36	**93.69**	**99.03**	99.14
				(0.28)	(0.29)	(0.09)	(0.09)			(0.24)	(0.24)	(0.09)	(0.09)
	CNN	100	96.76	**96.69**	91.63	99.17	**99.69**	100	94.16	**98.39**	82.50	97.38	**99.79**
				(0.32)	(0.31)	(0.10)	(0.10)			(0.33)	(0.28)	(0.12)	(0.11)
	No. of obs.	11,543		1,051		10,492		20,000		2,663		17,337	

tie-breaking. This reduces the macro-average F_1 score to 82% and 75% for the two datasets, respectively.

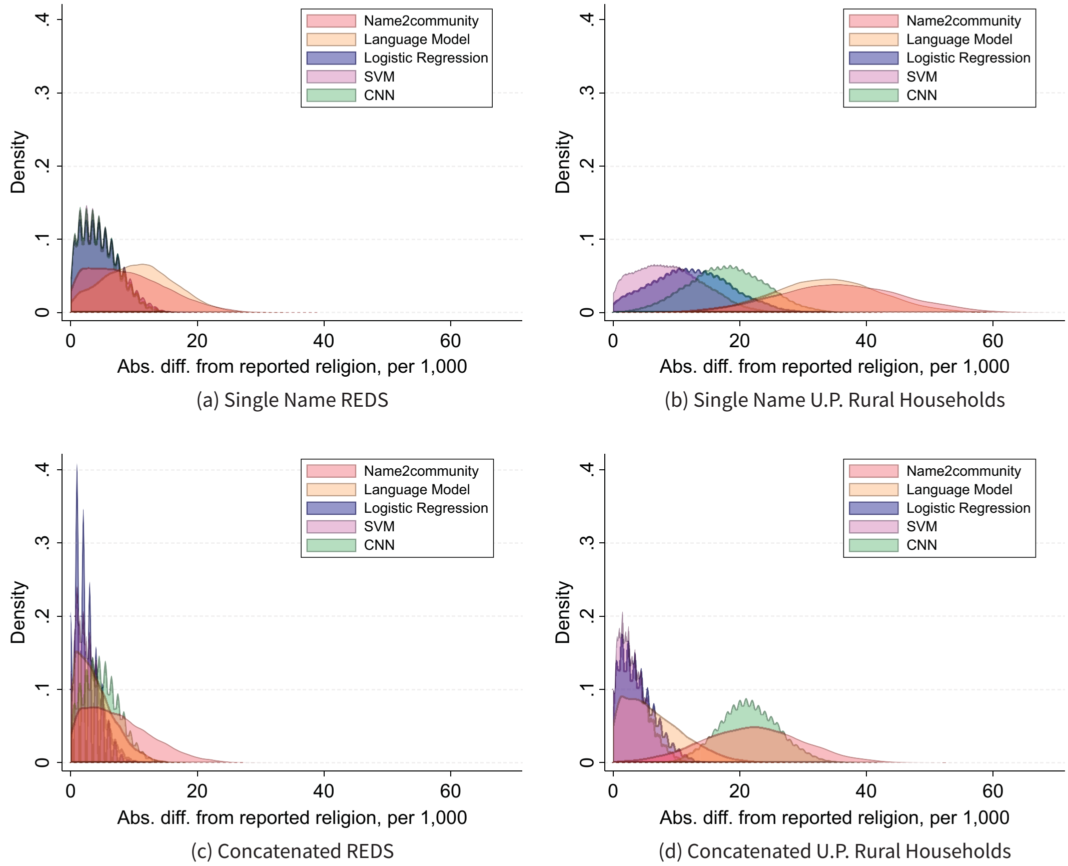
On the other hand, bag-of-*n*-grams models perform exceptionally well and have 100% coverage. Among the names that could be assigned a class using Name2community, the overall accuracy is significantly higher for LR, SVM, and CNN for both the test sets. In our experiments, Name2community was also orders of magnitude slower than character-based models. It could only predict 0.4 names/second. In contrast, LR, SVM, and CNN predicted 50,000–500,000 names/second. Therefore, predicting religion for the entire REDS test set comprising 11,543 names took over 8 hours using Name2community for single name model and nearly 13 hours with concatenated names. On the other hand, LR, SVM, and CNN only took 0.035–0.4 seconds to predict all the names. This makes Name2community less scalable to large datasets that may comprise millions of observations and further limits its viability as a general religion detector for South Asia. The coverage for language model is almost perfect, except in a few cases where the perplexity is infinity. It could predict approximately 2,500 names/second, which is slower than other character-based models but reasonable compared to Name2community. However, the performance of language model is worse for both the test sets with much lower F_1 scores. Comparing among the bag-of-*n*-grams models and CNN reveals that the accuracy of SVM and CNN is significantly higher than LR at the 5% level of significance. The performance of SVM and CNN is comparable. They also have higher macro-average F_1 scores than LR, though the difference is small.

We note that performance is lower on U.P. Rural Households data compared to REDS. This might be due to the difference in name length distributions across the two datasets.[24] To verify this, we restrict our sample to observations having only a first name representing 60% of the U.P. Rural Households data and 5% of the REDS test set. The F_1 score for LR, SVM, and CNN reduces to approximately 85% and is now comparable for both datasets.[25] For observations having both first and last names, the F_1 score is around 96% and is again comparable for the two datasets. This leads us to expect that including the name of a relative such as parent/spouse might improve the performance of the models—especially for the U.P. Rural Households data for which a majority of individuals only have information on the first name.

The results improve when we enrich our data by concatenating individual names with their parent/spouse's name (Panel B of Table 4). This is primarily driven by better identification of individuals for whom we only have the first name. The language model continues to perform worse than other character-based models. The F_1 score for CNN is now lower than LR and SVM. The overall accuracy is also lower for CNN with the difference being statistically significant at the 1% level. The recall for Muslim names is worse for CNN resulting in less balanced predictions. We also note that the coverage for Name2community increases by 7 percentage points for REDS and 15 percentage points for the U.P. Rural Households data. However, this is accompanied by only a marginal increase in the macro-average F_1 score. Thus, there are limited gains from providing richer data to Name2community.

Figure 2 shows the performance across models in predicting the aggregate Muslim share. For this, we follow the approach in Clark *et al.* (2021) and bootstrap 10,000 draws with a sample size of 1,000 per draw ($\approx$ population in a typical polling station in India). We then take the absolute difference between actual Muslim count and estimated Muslim count. We find that Name2community and language model almost always perform worse in estimating aggregate religious composition. In contrast, SVM and LR perform better and both have similar accuracy.

24 As discussed in Section 3, the average name length in REDS is around 15 characters, whereas for the U.P. Rural Households data, it is close to 9. This is explained by a majority of observations in U.P. Rural Households data having only first names.
25 F_1 score for language model reduces to 80% for these names. The coverage for Name2community reduces dramatically to only around 20%–30%.



(a) Single Name REDS
(b) Single Name U.P. Rural Households
(c) Concatenated REDS
(d) Concatenated U.P. Rural Households

Figure 2. Density plots of absolute difference in reported and estimated religious counts per 1,000 people.

The performance of CNN is comparable to LR and SVM for REDS. However, LR and SVM slightly outperform CNN for the U.P. Rural Households dataset. The differences in median accuracy between SVM and Name2community are 4.37 and 27.68 per 1,000 for REDS and U.P. Rural Households datasets, respectively, for single name models. The corresponding differences are 4.72 and 19.04 for concatenated models. These differences are sizeable and indicate that improved coverage and accuracy of the character-based models are meaningful even when estimating aggregate Muslim shares. We report the density plots of effective number of imputed religions, defined as the inverse of Herfindahl index, at the individual level in Figures 6 and 7 in Section E of the Supplementary Material. They show that the level of uncertainty for Name2community is higher than character-based models. The variation of error rates with effective number of imputed religions, shown in Figures 8 and 9 in Section E of the Supplementary Material, indicates that average error increases with uncertainty in model predictions.

One might expect bias against women who take up their husbands' names—especially in interfaith marriages. However, according to the 2019–2020 PEW Research Center survey less than 1% marriages in India are interfaith.[26] Furthermore, we restrict the sample to women for both the test sets and report the results in Table 10 in Section F of the Supplementary Material. The F_1 score is lower for single name models owing to a small fraction (10%) of women in the training data. However, LR, SVM, and CNN continue to outperform the baseline models. Therefore, character-based models have an advantage even when inferring the religion of women from names is of particular interest (as in Field, Jayachandran, and Pande 2010). The performance improves

26 In an earlier research, Goli, Singh, and Sekher (2013) use the nationally representative Indian Human Development Survey (2005) and find that only 2.21% women married outside their religion.

substantially when we include the name of parent/spouse—who usually share the same religion—and becomes comparable to that of the unrestricted sample.

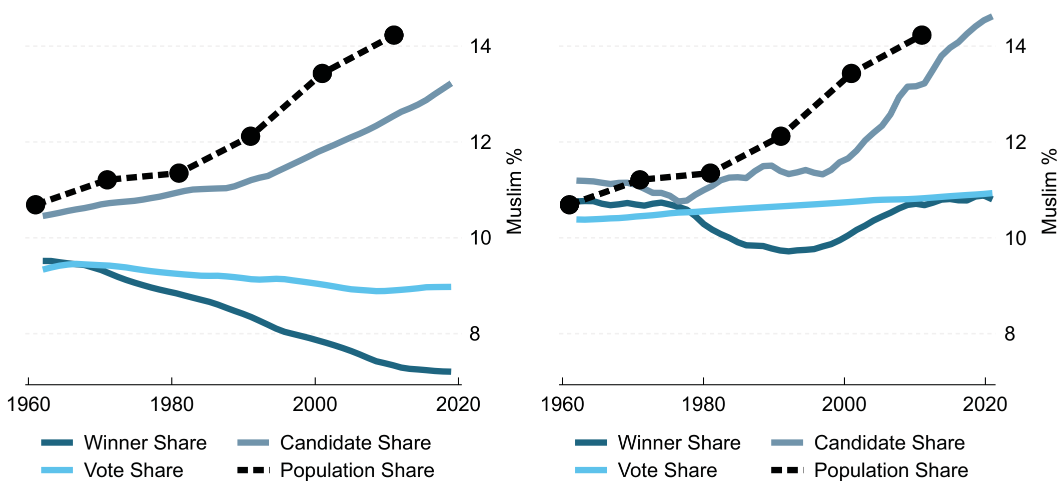
Summing up, character-based models perform better than Name2community and language model and can accurately infer religion at the individual/household level. Single name SVM and CNN perform slightly better than LR at the individual level. However, when estimating the aggregate composition, SVM and LR are comparable, while CNN performs slightly worse. Our preferred model is SVM given that CNN also requires extensive hyperparameter tuning, and therefore, might be harder to train. However, when probabilities are of interest, LR is preferable as SVM does not directly return probabilities. We also perform multi-religion classification. The character-based models still outperform Name2community across all the classes. See Section G of the Supplementary Material for a detailed discussion.

To illustrate the applicability of our methods in a different context, we also apply them to the task of race/ethnicity inference in the United States. We find that character-based models are substantially better than BISG at estimating the counts of Black voters who are the largest minority group, and therefore, more vulnerable to discrimination (Cikara, Fouka, and Tabellini 2022). This may be especially useful for guiding studies on Black voter behavior and political representation (Cascio and Washington 2014; Washington 2006). We discuss this in detail in Section A of the Supplementary Material.

7 Muslim Representation in India

In this section, we discuss an application of our work in the electoral context. We examine the temporal patterns of Muslim representation in the Indian legislature during 1962–2021. For this, we use the Indian Elections Dataset compiled by Agarwal *et al.* (2021). The dataset contains information on all the candidates and their electoral outcomes in the national- and state-level elections. The national elections data comprise 71,799 candidates in 8,277 races (4,524 unique winners), whereas the state-level data have 373,290 candidates in 54,143 races (34,924 unique winners). We classify candidate names using the binary SVM classifier. We exclude constituencies reserved for scheduled castes and scheduled tribes in our analysis.[27]

We report the results in Figure 3. Figure 3a shows the results for the national elections. We find that the Muslim candidate share increased commensurately with the Muslim population



(a) General Elections (b) Assembly Elections

Figure 3. Muslim representation in Indian Politics during 1962–2021.

27 We get virtually identical results if we use LR or CNN instead of SVM. The results are qualitatively similar if we include all the constituencies. Results are available on request.

share in India (based on the Census data). In contrast, the Muslim vote share stagnated and the share of Muslim legislators steadily declined. One possible explanation for this could be splitting of votes among different Muslim candidates. Figure 3b shows the aggregate results for state elections across India. Again, the share of Muslim candidates has increased with the overall Muslim population share. However, the vote share of Muslim contestants has remained almost constant. Muslims are underrepresented at both levels but the extent of under-representation is lower in state assemblies than at the national level. This could be explained by a combination of a higher Muslim vote share and a more proportionate mapping of vote share into seat share due to the smaller size of assembly constituencies in comparison to parliamentary constituencies.

8 Ethical Implications

Religion is a sensitive issue and religious minorities face unfair treatment in many parts of the world. For example, the 2004 French headscarf ban adversely affected the educational attainment and labor market outcomes of Muslim women (Abdelgadir and Fouka 2020). Characterized by strong communal politics, India is no exception to this. India's Independence from colonial rule was accompanied by large-scale violence between Hindus and Muslims, and the underlying social tensions have persisted to date (Jha 2013; Varshney and Wilkinson 2016). Politicians have frequently weaponized these tensions for electoral gains (Wilkinson 2006). There is evidence of religious discrimination beyond politics too, for instance, in urban housing rental markets (Thorat *et al.* 2015) and labor markets (Thorat and Attewell 2007). Therefore, it is of utmost importance to include a discussion on ethical implications of our work.

A potential risk that comes with any new technology is that it may be used by nefarious actors. In our case, one could argue that the inferences made by our models could exacerbate religious discrimination and violence. This might also pose a risk to people incorrectly classified among the targeted group. However, as discussed, religious connotations of names are well known in India. This means that while making individual level decisions such as hiring or renting property, employers, and landlords do not require an algorithm to identify religion. Likewise, local rioters can easily use publicly available name-lists to manually identify religion with near perfect accuracy. Moreover, government actors already have access to detailed information on religion through the Indian Census and other sources. Therefore, the algorithm is less relevant for individual actors or the government.

On the other hand, by making data more accessible to researchers, we expect our work will help highlight any discrimination and deprivations experienced by minorities. This could, in turn, help formulate policies to protect vulnerable groups. Therefore, we expect the benefits of such a technology to outweigh the potential harms. We caution that names may not strictly correspond to religious or other identity groups. Therefore, our classification exercise should be interpreted as a probabilistic rather than a rigid mapping.

9 Conclusion

We infer religion from South Asian names using character-based machine-learning models. While dictionary-based methods suffer from slow speed, low coverage, and low accuracy, character-based models perform exceptionally well and can easily be scaled to massive datasets. Given the lack of fine-grained data on religion and abundance of public name lists, this is a viable way to generate individual level data on group identity and also obtain precise group composition estimates. We apply our methods to the question of Muslim representation in India and observe a declining trend despite a rise in their population share. We use LRP to explain the predictions of one of our models and detect linguistic differences in names across religions. More than 40 years back, Emeneau (1978) observed that exhaustive studies on South Asian onomastics could only be undertaken with the aid of sophisticated computer programs due to the sheer size of the population. We hope that our analysis motivates more in-depth linguistic studies.

More broadly, this work can serve as a methodological guide outside South Asia to identify groups for whom the underlying names might have different linguistic roots. For example, in Latin America, White names reflect the predominant language (Spanish) spoken by them while the indigenous names strongly reflect indigenous languages (Guarani, Quechua, Nahuatl, and Aymara, etc.). Similarly, in Nigeria, Hausa names belonging to the Afro-Asiatic family are distinct from Yoruba and Igbo names having roots in Niger-Congo linguistic family. We also demonstrate that our models can aid the existing BISG packages in the United States—replacing surname dictionaries for calculating prior probabilities of an individual's race/ethnicity and reduce bias in estimating Black voter count. Thus, our work can contribute toward a richer understanding of economic conditions of various groups, discrimination, residential segregation, conflict, and identity politics; it can also facilitate research on group behavior on social media platforms.

Acknowledgments
We are indebted to Mithilesh Chaturvedi, Sabyasachi Das, Sudha Rao, and Feyaad Allie for helpful feedback and suggestions. We also thank the editorial team at Political Analysis and four anonymous reviewers for their detailed engagement and suggestions toward the improvement of the manuscript.

Data Availability Statement
The replication code for this article is available at https://doi.org/10.7910/DVN/JOEVPN (Chaturvedi and Chaturvedi 2023). Instructions for accessing Rural Economic & Demographic Survey (REDS) data are available at http://adfdell.pstc.brown.edu/arisreds_data/readme.txt.

Conflict of Interest
The authors report no conflict of interest.

Supplementary Material
For supplementary material accompanying this paper, please visit https://doi.org/10.1017/pan.2023.6.

References
Abdelgadir, A., and V. Fouka. 2020. "Political Secularism and Muslim Integration in the West: Assessing the Effects of the French Headscarf Ban." *American Political Science Review* 114 (3): 707–723.

Adida, C. L., D. D. Laitin, and M.-A. Valfort. 2010. "Identifying Barriers to Muslim Integration in France." *Proceedings of the National Academy of Sciences* 107 (52): 22384–22390.

Agarwal, A., et al. 2021. "TCPD Indian Elections Data v2.0." Technical report, Trivedi Centre for Political Data, Ashoka University.

Algan, Y., C. Malgouyres, T. Mayer, and M. Thoenig. 2022. "The Economic Incentives of Cultural Transmission: Spatial Evidence from Naming Patterns across France." *The Economic Journal* 132 (642): 437–470.

Ambekar, A., C. Ward, J. Mohammed, S. Male, and S. Skiena. 2009. "Name-Ethnicity Classification from Open Sources." In *Proceedings of the 15th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 49–58. New York, NY: ACM.

Ananat, E. O., and E. Washington. 2009. "Segregation and Black Political Efficacy." *Journal of Public Economics* 93 (5–6): 807–822.

Ancona, M., E. Ceolini, C. Öztireli, and M. Gross. 2018. "Towards Better Understanding of Gradient-Based Attribution Methods for Deep Neural Networks." In *6th International Conference on Learning Representations (ICLR)*, 1–16. Vancouver; arXiv:1711.06104.

Arras, L., F. Horn, G. Montavon, K.-R. Müller, and W. Samek. 2017. "What Is Relevant in a Text Document?": An Interpretable Machine Learning Approach." *PLoS One* 12 (8): e0181142.

Bach, S., A. Binder, G. Montavon, F. Klauschen, K.-R. Müller, and W. Samek. 2015. "On Pixel-Wise Explanations for Non-linear Classifier Decisions by Layer-Wise Relevance Propagation." *PLoS One* 10 (7): e0130140.

Bertrand, M., and S. Mullainathan. 2004. "Are Emily and Greg more Employable than Lakisha and Jamal? A Field Experiment on Labor Market Discrimination." *American Economic Review* 94 (4): 991–1013.

PA

Bhalotra, S., I. Clots-Figueras, G. Cassan, and L. Iyer. 2014. "Religion, Politician Identity and Development Outcomes: Evidence from India." *Journal of Economic Behavior & Organization* 104: 4–17.

Bhalotra, S., I. Clots-Figueras, L. Iyer, and J. Vecci. 2021. "Leader Identity and Coordination." *The Review of Economics and Statistics* 105: 175–189.

Bhalotra, S., C. Valente, and A. Van Soest. 2010. "The Puzzle of Muslim Advantage in Child Survival in India." *Journal of Health Economics* 29 (2): 191–204.

Biavaschi, C., C. Giulietti, and Z. Siddique. 2017. "The Economic Payoff of Name Americanization." *Journal of Labor Economics* 35 (4): 1089–1116.

Cascio, E. U., and E. Washington. 2014. "Valuing the Vote: The Redistribution of Voting Rights and State Funds Following the Voting Rights Act of 1965." *The Quarterly Journal of Economics* 129 (1): 379–433.

Cavnar, W. B., and J. M. Trenkle. 1994. "N-Gram-Based Text Categorization." In *Proceedings of SDAIR-94, 3rd Annual Symposium on Document Analysis and Information Retrieval*, 161–175. Las Vegas, NV.

Chang, J., I. Rosenn, L. Backstrom, and C. Marlow. 2010. "Epluribus: Ethnicity on Social Networks." In *Fourth International AAAI Conference on Weblogs and Social Media*. Palo Alto, CA: AAAI.

Chaturvedi, R., and S. Chaturvedi. 2023. "Replication Data for: It's All in the Name: A Character Based Approach to Infer Religion." https://doi.org/10.7910/DVN/JOEVPN.

Chaturvedi, S., S. Das, and K. Mahajan. 2021. "The Importance of Being Earnest: What Drives the Gender Quota Effect in Politics?" Available at *SSRN* 3962068.

Chhibber, P. K., and S. Shastri. 2014. *Religious Practice and Democracy in India*. New York: Cambridge University Press.

Cikara, M., V. Fouka, and M. Tabellini. 2022. "Hate Crime Towards Minoritized Groups Increases as They Increase in Sized-Based Rank." *Nature Human Behaviour* 6 (11): 1537–1544. *Hate Crime Increases with Minoritized Group Rank*.

Clark, J. T., J. A. Curiel, and T. S. Steelman. 2021. "Minmaxing of Bayesian Improved Surname Geocoding and Geography Level Ups in Predicting Race." *Political Analysis* 30: 456–462.

Collobert, R., J. Weston, L. Bottou, M. Karlen, K. Kavukcuoglu, and P. Kuksa. 2011. "Natural Language Processing (Almost) from Scratch." *Journal of Machine Learning Research* 12: 2493–2537.

Cutler, D. M., E. L. Glaeser, and J. L. Vigdor. 1999. "The Rise and Decline of the American Ghetto." *Journal of Political Economy* 107 (3): 455–506.

Elliott, M. N., P. A. Morrison, A. Fremont, D. F. McCaffrey, P. Pantoja, and N. Lurie. 2009. "Using the Census Bureau's Surname List to Improve Estimates of Race/Ethnicity and Associated Disparities." *Health Services and Outcomes Research Methodology* 9 (2): 69–83.

Emeneau, M. B. 1978. "Towards an Onomastics of South Asia." *Journal of the American Oriental Society* 98 (2): 113–130.

Field, E., S. Jayachandran, and R. Pande. 2010. "Do Traditional Institutions Constrain Female Entrepreneurship? A Field Experiment on Business Training in India." *American Economic Review* 100 (2): 125–129.

Field, E., M. Levinson, R. Pande, and S. Visaria. 2008. "Segregation, Rent Control, and Riots: The Economics of Religious Conflict in an Indian City." *American Economic Review* 98 (2): 505–510.

Fouka, V. 2019. "How Do Immigrants Respond to Discrimination? The Case of Germans in the US during World War I." *American Political Science Review* 113 (2): 405–422.

Fouka, V. 2020. "Backlash: The Unintended Effects of Language Prohibition in US Schools after World War I." *The Review of Economic Studies* 87 (1): 204–239.

FryerJr , R. G., and S. D. Levitt. 2004. "The Causes and Consequences of Distinctively Black Names." *The Quarterly Journal of Economics* 119 (3): 767–805.

Gaikwad, N., and G. Nellis. 2017. "The Majority-Minority Divide in Attitudes toward Internal Migration: Evidence from Mumbai." *American Journal of Political Science* 61 (2): 456–472.

Goli, S., D. Singh, and T. Sekher. 2013. "Exploring the Myth of Mixed Marriages in India: Evidence from a Nation-Wide Survey." *Journal of Comparative Family Studies* 44 (2): 193–206.

Guiso, L., P. Sapienza, and L. Zingales. 2003. "People's Opium? Religion and Economic Attitudes." *Journal of Monetary Economics* 50 (1): 225–282.

Harris, J. A. 2015. "What's in a Name? A method for Extracting Information about Ethnicity Fromnames." *Political Analysis* 23 (2): 212–224.

Heath, O., G. Verniers, and S. Kumar. 2015. "Do Muslim Voters Prefer Muslim Candidates? Co-Religiosity and Voting Behaviour in India." *Electoral Studies* 38: 10–18.

Imai, K., and K. Khanna. 2016. "Improving Ecological Inference by Predicting Individual Ethnicity from Voter Registration Records." *Political Analysis* 24 (2): 263–272.

Iyer, S. 2016. "The New Economics of Religion." *Journal of Economic Literature* 54 (2): 395–441.

Iyer, S. 2018. *The Economics of Religion in India*. Cambridge, MA: Harvard University Press.

Jauhiainen, T., K. Lindén, and H. Jauhiainen. 2017. "Evaluation of Language Identification Methods Using 285 Languages." In *Proceedings of the 21st Nordic Conference on Computational Linguistics*, 183–191. Gothenburg, Sweden: Association for Computational Linguistics.

Jauhiainen, T., K. Lindén, and H. Jauhiainen. 2019a. "Language Model Adaptation for Language and Dialect Identification of Text." *Natural Language Engineering* 25 (5): 561–583.

Jauhiainen, T., M. Lui, M. Zampieri, T. Baldwin, and K. Lindén. 2019b. “Automatic Language Identification in Texts: A Survey.” *Journal of Artificial Intelligence Research* 65: 675–782.
Jensen, J. L., D. Karell, C. Tanigawa-Lau, N. Habash, M. Oudah, and D. Fairus Shofia Fani. 2021. “Language Models in Sociological Research: An Application to Classifying Large Administrative Data and Measuring Religiosity.” *Sociological Methodology* 52: 00811750211053370.
Jha, S. 2013. “Trade, Institutions, and Ethnic Tolerance: Evidence from South Asia.” *American Political Science Review* 107 (4): 806–832.
Knowles, R., J. Carroll, and M. Dredze. 2016. “Demographer: Extremely Simple Name Demographics.” In *Proceedings of the First Workshop on NLP and Computational Social Science*, 108–113. Austin, TX: Association for Computational Linguistics.
Kundu, A. 2014. “Post Sachar evaluation committee report.” In *Ministry of Minority Affairs, Government of India, New Delhi*, 1950–1995.
LeCun, Y., et al. 1989. “Backpropagation Applied to Handwritten Zip Code Recognition.” *Neural Computation* 1 (4): 541–551.
Lee, J., H. Kim, M. Ko, D. Choi, J. Choi, and J. Kang. 2017. “Name Nationality Classification with Recurrent Neural Networks.” In *IJCAI*, 2081–2087. Palo Alto, CA: AAAI Press.
Mateos, P. 2007. “A Review of Name-Based Ethnicity Classification Methods and their Potential in Population Studies.” *Population, Space and Place* 13 (4): 243–263.
Metcalf, B. D. 2009. *Islam in South Asia in Practice*, Vol. 33. Princeton: Princeton University Press.
Mitra, A., and D. Ray. 2014. “Implications of an Economic Theory of Conflict: Hindu-Muslim Violence in India.” *Journal of Political Economy* 122 (4): 719–765.
Müller-Crepon, C., and P. Hunziker. 2018. “New Spatial Data on Ethnicity: Introducing SIDE.” *Journal of Peace Research* 55 (5): 687–698.
Nellis, G., et al. 2016. “Do Parties Matter for Ethnic Violence? Evidence from India.” *Quarterly Journal of Political Science* 11 (3): 249–277.
Oreopoulos, P. 2011. “Why Do Skilled Immigrants Struggle in the Labor Market? A Field Experiment with Thirteen Thousand Resumes.” *American Economic Journal: Economic Policy* 3 (4): 148–171.
Sachar, R., et al. 2006. “Social, Economic and Educational Status of the Muslim Community of India.” Technical report, East Asian Bureau of Economic Research.
Sahgal, N., J. Evans, A. Salazar, K. Starr, and M. Corichi. 2021. “Religion in India: Tolerance and Segregation.” Technical report.
Schimmel, A. 1997. *Islamic Names*. Edinburgh: Edinburgh University Press.
Susewind, R. 2015. “What’s in a Name? Probabilistic Inference of Religious Community from South Asian Names.” *Field Methods* 27 (4): 319–332.
Thorat, S., and P. Attewell. 2007. “The Legacy of Social Exclusion: A Correspondence Study of Job Discrimination in India.” *Economic and Political Weekly* 42: 4141–4145.
Thorat, S., A. Banerjee, V. K. Mishra, and F. Rizvi. 2015. “Urban Rental Housing Market.” *Economic & Political Weekly* 27: 47–53.
Torres, M., and F. Cantú. 2022. “Learning to See: Convolutional Neural Networks for the Analysis of Social Science Data.” *Political Analysis* 30 (1): 113–131.
Torvik, V. I., and S. Agarwal. 2016. “Ethnea–An Instance-Based Ethnicity Classifier Based on Geo-Coded Author Names in a Large-Scale Bibliographic Database.” In *International Symposium on Science of Science, Washington DC, USA*.
Treeratpituk, P., and C. L. Giles. 2012. “Name-Ethnicity Classification and Ethnicity-Sensitive Name Matching.” In *Twenty-Sixth AAAI Conference on Artificial Intelligence*. Palo Alto, CA: AAAI Press.
Varshney, A., and S. Wilkinson. 2016. “Varshney-Wilkinson Dataset on Hindu-Muslim Violence in India, 1950–1995, Version 2.” Inter-university Consortium for Political and Social Research [distributor], 2006-02-17. https://doi.org/10.3886/ICPSR04342.
Vatanen, T., J. Väyrynen, and S. Virpioja. 2010. “Language Identification of Short Text Segments with N-Gram Models.” In *Proceedings of the Seventh International Conference on Language Resources and Evaluation (LREC’10)*. Valletta, Malta: European Language Resources Association (ELRA).
Washington, E. 2006. “How Black Candidates Affect Voter Turnout.” *The Quarterly Journal of Economics* 121 (3): 973–998.
Wilkinson, S. I. 2006. *Votes and Violence: Electoral Competition and Ethnic Riots in India*. Cambridge: Cambridge University Press.
Wood-Doughty, Z., N. Andrews, R. Marvin, and M. Dredze. 2018. “Predicting Twitter User Demographics from Names Alone.” In *Proceedings of the Second Workshop on Computational Modeling of People’s Opinions, Personality, and Emotions in Social Media*, 105–111. New Orleans, LA: Association for Computational Linguistics.
Zhang, X., J. Zhao, and Y. LeCun. 2015. “Character-Level Convolutional Networks for Text Classification.” In *Advances in Neural Information Processing Systems*, 649–657. Red Hook, NY: Curran Associates, Inc.