

Intrinsic race-telltales from clinical notes? Identification of bias and its interpretation

Hanyue Zhou¹ and Dan Ruan¹²

¹Department of Bioengineering, University of California, Los Angeles, California, U.S

²Department of Radiation Oncology, University of California, Los Angeles, California, U.S



INTRODUCTION

Healthcare disparities are the differences in healthcare quality, access and outcomes adversely affecting members of racial minority groups and other social disadvantaged populations [1]. It is necessary to understand the source of disparities and identify the contributing factors, which in return will provide effective interventions to reduce the racial and ethnic disparities in health care.

Classic disparity analysis follows the general route of establishing a standing hypothesis by specifying study covariates and endpoints, collecting the relevant data, and performing statistical tests. However, both medial comorbidities and implicit bias can manifest in a wide variety of manners and it is almost impossible to designate comprehensive covariate signatures for conventional hypothesis testing framework. In this project, we focus on race-related analysis and propose novel language processing methods on medical notes to study the intrinsic differentiation between the black and white sub-cohorts .

AIM

In the era of precision medicine, while it is desirable to personalize treatment according to individual patient characteristics, it is important to take caution against differential treatment caused by implicit bias from caregivers. Therefore, it is important to (1) identify the existence of bias and (2) interpret the prominent contributing factors.

METHOD

Preprocessing:

- Note de-identified (ID numbers, names, dates and times removal) by pattern matching approaches
- Punctuation removed
- Word embeddings pretrained by word2vec model [2]

Baseline model:

- Document embeddings calculated as the average of word and ngram embeddings, based on Fasttext [3]
- Variational information bottleneck (VIB) [4][5] as regularization
- Linear support vector machine (SVM) to classify document embeddings
- Words and ngrams with big projections on the normal direction of SVM hyperplane are important

Hierarchical attention networks (HAN):

- Bidirectional gated recurrent unit (GRU) modules and attention layers on word level and sentence level [6]
- Variational information bottleneck (VIB) as regularization on word level and sentence level
- Words and sentences with big attention weights are important

RESULTS

Under IRB approved protocol, clinical notes were collected for PC treatment. We performed binary classification between 71 reports from black patients and 71 reports from white patients.

We illustrate the classification performance and the identified important text features by each model.

Accuracy/ Precision/ Recall	Baseline Model	HAN
Training	0.83/ 0.83/ 0.83	1/1/1
Validation	0.83/ 0.84/ 0.83	0.94/ 0.95/ 0.94
Testing	0.65/ 0.66/ 0.65	0.90/ 0.92/ 0.90

Classification results by the two models

	Words	Sentences
Black	4+3	cancer t1c ipsa 8.7 bgs 4+3 3+4 3+3 in cores no pni s/p
	gleason	imrt without adt completed in april 2014
	3+4	
	bgs	
	1	male with unfavorable intermediate-risk prostate
	nocturia	cancer t1c ipsa 8.7 bgs 4+3 3+4 3+3 in cores no pni s/p
	cancer	imrt without adt completed in april 2014
	3+3	
	4+5	male with a history of pt2cno prostate cancer gleason
	constipation	3+4 ipsa 7.1 s/p prostatectomy on with positive
	baseline	margins at the left and right base focal pni
	taking	he is doing well urinary frequency issues resolved
	x	baseline constipation nocturia x 1 not on flomax
	4+4	
	8.7	he is doing well urinary frequency issues resolved bm
White	pathology	normal no brbpr nocturia x 1 not on flomax
	behavior	
	ipsa	male with unfavorable intermediate-risk prostate
	not	cancer ct2c bgs 4+3 ipsa 4 s/p definitive imrt in with
	cores	one year of adt
	ct2c	
	radiology	ergocalciferol units capsule take 1 capsule 50,000 units
	13.5	total by mouth every seven 7 day
	capsule	
	questionnair	ct2 gs 3+4x4 3+5x1 4+3x1 and 3+3x3 total cores pni+
	e	s/p sbrrt on high-risk protocol with 9 months adt
	3+3x3	
	total	
	systems	cores involved bgs 4+3x1 3+3x1 pni+ radiation history
	urgency	40gy sbrrt prostate and 25gy pelvic nodes with 9
	4+3x1	months adt on high-risk study completed on data of
	cores	interval history prostate-fu data fields data comments
	gs	most recent psa level 1.3 t 267 recent psa date
	4+3	m/d/yyyy biochemical failure bf no if bf date m/d/yyyy
	3+3x1	local failure lf no
	pni+	biopsy showed positive cores pni+ gs 4+4x1 4+3x5
	suggestion	3+4x3 and 3+3x1
	capsular	
	copies	doing well overall very satisfied stable urinary issues
	involvement	nocturia x 1 bm normal no brbpr
	3+4x4	
	description	some urinary urgency improving but admits to drinking
	bleeding	some 10 of tea a day
	taking	4+3 involving cores 10-20 of each core s/p adt 3 month
	mri	and hdr brachytherapy 38 gy in 4 fx completed
	limited	
	imaging	capsular margin suggestion of capsular involvement as
	4+4x1	described above
	vitamin	

Most important words and sentences identified by the VIB regularized HAN model; importance measured by attention weights

	Words and Phrases	Projection	Word and Phrase Contexts
Black	hh	1.467	PSA, Total Latest Ref Range: 0-4.5
	periatrrial	1.380	ng/mL 41.3 (HH) 10:00 AM
	0-4.5	1.304	subsequent diagnosis of brain
	0-4.5 ng/ml	1.288	metastases in 6/2013, s/p SRS to 2
	ng/ml	1.283	lesions completed, complicated by
	range 0-4.5	1.261	the development of radiation
	ng/ml		necrosis associated with the treated L
	hh pm	1.235	periatrrial lesion
	hh am	1.230	
	range	1.224	Liver: A few scattered subcentimeter
	l periatrrial	1.219	low density lesions are too small to
	density	1.214	characterize but stable.
	range ng/ml	1.206	
	261	1.175	Testosterone Latest Ref Range: 200 -
	range 200-1000	1.170	1,000 ng/dL 261 141 (L)
White	ng/ml 41.3 hh	1.168	post-op PSA 0.43
	0.43	1.161	pT1N0 papillary thyroid
	pt1n0	1.151	microcarcinoma
	indicative	1.149	The patient's elevated PSA in the
	range 0-3.5	1.140	setting of prior surgery is likely
	ng/ml		indicative of a biochemical
	0.71	1.135	recurrence.
	questionnaire	-1.793	A 14 point review of systems
	systems	-1.661	questionnaire was given to the
	questionnaire		patient and reviewed by the
	copies	-1.648	attending physician. Signed copies
	adl	-1.632	are available in the chart.
	instrumental adl	-1.589	Pain: Grade 2: Moderate pain;
	frequency/rigidi	-1.574	ty limiting instrumental ADL
	systems	-1.570	Erectile Dysfunction: Grade 2:
	instrumental	-1.549	Decrease in erectile function
	shim	-1.518	(frequency/rigidity of erections)
	anorgasmia	-1.469	QOL score: Delighted SHIM score:
	thromboemboli	-1.463	0: no attempt since procedure
	sm		
	adversely	-1.461	Anorgasmia: Grade 0
	hyperpigmentat	-1.450	The acute and chronic risks of
	ion		anesthesia and brachytherapy
	limiting	-1.447	(including but not limited to
	instrumental adl		bleeding, infection,
	1+2	-1.441	thromboembolism, organ obstruction
	urgency/inconti	-1.430	or dysfunction) were described.
	nence		
	delighted shim	-1.421	decrease in sexual interest not
	meds shim	-1.402	adversely affecting relationship
	instrumental adl	-1.393	Skin Hyperpigmentation: Grade 0
	urinary		
	rigidity	-1.391	Biopsy Gleason Score Total (1+2) 9

Most important words and phrases identified by the baseline model; importance measured by projection component

datetime	note	id	radiation	oncology	note	prostate	filed
office	visit	note	patient		prostate		follow-up
mrm	date	mrm	of	service	doi		
referring	practitioner	primary	care	provider	no	primary	
care	provider	primary	care	provider	resident	physician	
attending	physician	chief	identifying	data/oncologic	history	2	year
is	prostate	menecarcinoma	2007	with	definitive	radiation	and
75	gy	lapron	imrt	adt	prostate	rt	
2009	prostate-fu	data	recent	psa	biochemical	failure	bf
comments	no	date	failure	biochemical	failure	bf	
bf	local	failure	m/d/yyyy	if	if	if	
local	failure	m/d/yyyy	if	if	if	if	
any	failure	within	rt	metastases	field	no	
complications	ctcae	interval	dr	history	0	patient	was
by	time	he	reported	addition	he	had	bothering
urinary	incontinence	he	leakage	occasional	he	was	not
recommended	minor	he	go	incontinence	he	was	not
and	minor	he	go	incontinence	he	was	not
well	nourished	he	go	incontinence	he	was	not
not	clinically	assess	lab	frequency	and	recommended	minor
consultation	for	the	minutes	of	the	patient	spant
with	the	patient	spant	in	detailed	question	team
session	cc	patient	care	team	radiation	oncology	
urology	radiation	oncology	radiation	oncology	am	note	deleted

An example report from a black patient: dark blocks represent tokens (punctuation, dates, times and IDs) that are redacted during preprocessing for HIPPA; attention weights are assigned to words (blues) and sentences (reds); the report reads from left to right and from top to bottom; the word weights are multiplied by the square of sentence weights

datetime	note	id	radiation	oncology	note	prostate	filed
office	visit	note	patient		prostate		follow-up
mrm	date	mrm	of	service	doi		
referring	practitioner	primary	care	provider	no	primary	
care	provider	primary	care	provider	resident	physician	
attending	physician	chief	identifying	data/oncologic	history	2	year
is	prostate	menecarcinoma	2007	with	definitive	radiation	and
75	gy	lapron	imrt	adt	prostate	rt	
2009	prostate-fu	data	recent	psa	biochemical	failure	bf
comments	no	date	failure	biochemical	failure	bf	
bf	local	failure	m/d/yyyy	if	if	if	
local	failure	m/d/yyyy	if	if	if	if	
any	failure	within	rt	metastases	field	no	
complications	ctcae	interval	dr	history	0	patient	was
by	time	he	reported	addition	he	had	bothering
urinary	incontinence	he	leakage	occasional	he	was	not
recommended	minor	he	go	incontinence	he	was	not
and	minor	he	go	incontinence	he	was	not
well	nourished	he	go	incontinence	he	was	not
not	clinically	assess	lab	frequency	and	recommended	minor
consultation	for	the	minutes	of	the	patient	spant
with	the	patient	spant	in	detailed	question	team
session	cc	patient	care	team	radiation	oncology	
urology	radiation	oncology	radiation	oncology	am	note	deleted

An example report from a white patient: dark blocks represent tokens (punctuation, dates, times and IDs) that are redacted during preprocessing for HIPPA; attention weights are assigned to words (blues) and sentences (reds); the report reads from left to right and from top to bottom; the word weights are multiplied by the square of sentence weights

REFERENCES

1. "CDC (Cent. Dis. Control Prev.) Minority health: definitions. Updated Feb. 2014, CDC, Atlanta,," 2014. [Online]. Available: <http://www.cdc.gov/minorityhealth/populations/REMP/definitions.htm>.
2. T. Mikolov, K. Chen, G. Corrado, and J. Dean, "Efficient estimation of word representations in vector space," *1st Int. Conf. Learn. Represent. ICLR 2013 - Work*.
3. A. Joulin, E. Grave, P. Bojanowski, and T. Mikolov, "Bag of Tricks for Efficient Text Classification," 2016.
4. A. A. Alemi, I. Fischer, J. V. Dillon, and K. Murphy, "Deep Variational Information Bottleneck," *5th Int. Conf. Learn. Represent. ICLR 2017 - Conf. Track Proc.*, Dec. 2016.
5. X. L. Li and J. Eisner, "Specializing Word Embeddings (for Parsing) by Information Bottleneck," pp. 2744–2754, 2019.
6. Z. Yang, D. Yang, C. Dyer, X. He, A. Smola, and E. Hovy, "Hierarchical attention networks for document classification," in *2016 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language*

ACKNOWLEDGEMENTS

The authors thank Varian for their financial support in building the information infrastructure for this study. They also acknowledge the department of radiation oncology, the ISS team at UCLA for their continuous support and consultations.

CONTACT INFORMATION

Hanyue (Nicole) Zhou: hyuezhou@gmail.com, Department of Bioengineering, UCLA