

## **Supplementary Materials for “Developing an automatic pipeline for analyzing chatter about health services from social media: A case study for Medicaid”**

### **Authors:**

\*Yuan-Chi Yang, PhD<sup>1</sup>,  
Mohammed Ali Al-Garadi, PhD<sup>1</sup>,  
Jane M. Zhu, MD, MPP, MSHP<sup>2</sup>,  
David Grande, MD, MPA<sup>3</sup>,  
Whitney Hogg-Bremer, BS<sup>1</sup>  
Abeed Sarker, PhD<sup>1</sup>

<sup>1</sup>Department of Biomedical Informatics, School of Medicine, Emory University, Atlanta, GA, United States;

<sup>2</sup> Division of General Internal Medicine and Geriatrics, Oregon Health & Science University, Portland, OR;

<sup>3</sup> Division of General Internal Medicine and Geriatrics, School of Medicine, University of Pennsylvania, Philadelphia, PA;

\*Corresponding author

Email: [yuan-chi.yang@emory.edu](mailto:yuan-chi.yang@emory.edu)

**Table 1:** List of keywords used for collecting Corpus 1

Type of keywords	keywords
Medicaid and its common misspellings	medicaid medicade medcaid medicad mediaid mediciad medicaide
Branded Names of Medicaid Agencies or Managed Care Organizations	firstcare excellusbcb mohealthnet badgercare wellsense medquest denalicare forwardhealth medi-cal mainecare tenncare kancare alamedaalliance hapmichigan ghcsw mywellsense cencalhealth mclarenhlthplan bhpartnershipma vapremierhealth

**Table 2:** List of handles used for collecting Corpus 2

Medicaid Agency	AHCCCSgov, IAHealthLink, MDMedicaid, MassHealth, MSMedicaid, OhioMedicaid, oksoonercare, CCOOregon, scmedicaid, TennCareRep, DVHAVermont, WA_Health_Care, HUSKYHealthCT, GADCH, CHCPF
Managed Care Organizations	PSHP, mercycareaz, YamhillCCO, PassportHealthP, SFHealthPlan, AmidaCareNY, AetnaMedicaidNC, magnolia_health, HNEinc, AmeriHealthLA, NetworkHealthWI, UPMCHealthPlan, Medi_CalHCO, ExcellusBCBS, LACareHealth, QuartzBenefits, CoordinatedCare, Buckeye_Health, MVPHealthCare, PartnershipHP, hampmichigan, HPofSanJoaquin, CHCTexas, HSCHealthCare, AnthemBC_News, SunHealthFL, SilverSummitHP, FidelisCare, PriorityHealth, Trilliumchp, HorizonNJHealth, LAHealthConnect, PresHealth, BCBSNM, ECCooperative, HPPlans, RMHP, MHS_Wisconsin, NHPRIHealth, bhpartnershipma, HealthShareOR, VAPremierHealth, IlliniCare, GHCSCW, ucaremn, CFHealthPlans, iCareHealthPlan, _KFHC, MDwiseInc, HSCSN_Inc, CalOptima, SunflowerHPlan, HealthfirstNY, TRUSTEDHP, NextLevelIL, affinityplan, CencalHealth, AbetterGateway, HealthChoiceUT, AmbetterTX, MyWellSense, AlohaCareHawaii, YourCare_Health, CrystalRunHP, HealthChoiceAZ, firstcare, CoCoHealth, BCBSTX, MHSIndiana, UPHealthPlan, MetroPlusHealth, NHhealthy, CDPHP, AskHMSA, myCHPW, McLarenHlthPlan, CAHandW, IEHP_healthcare, BlueCrossMN, setonfamily, BMCHealthNet

**Table 3:** Classification performances of the classifiers on the test sets of Corpus-1 & 2.

Classification Algorithm	test set (Corpus-1)	Precision (.XX)						Recall (.XX)					
		a	c	i	n	p	o	a	c	i	n	p	o
	NB	17	10	16	54	88	33	9	57	49	60	59	36
	SVM	46	<b>43</b>	37	73	88	75	31	<b>70</b>	47	77	90	34
	RF	50	60	41	73	81	84	3	24	14	76	94	32
	KNN	0	35	16	48	89	17	0	22	6	65	53	64
	NN	21	36	19	70	87	46	14	46	29	71	84	46
	BLSTM	0	50	34	71	85	78	0	57	29	79	91	39
	BERT	59	<b>55</b>	46	79	92	78	49	<b>73</b>	61	86	91	46
	test set (Corpus-2)	Precision (.XX)						Recall (.XX)					
		a	c	i	n	p	o	a	c	i	n	p	o
	NB	0	67	14	15	0	60	0	82	47	33	0	11
	SVM	100	<b>79</b>	64	43	0	78	33	<b>91</b>	41	50	0	78
	RF	0	79	0	67	8	67	0	83	0	33	14	78
KNN	0	92	0	0	0	43	0	13	0	0	0	100	
NN	100	79	43	25	4	75	33	70	53	17	14	68	
LSTM	0	79	57	71	14	76	0	75	24	83	71	65	
BERT	100	<b>92</b>	93	40	33	93	67	<b>92</b>	76	67	57	87	

**Table 4:** The 10 highest-ranking TFIDF bigrams and trigrams in each class

<b>Class</b>	<b>bigrams</b>	<b>trigrams</b>
a	<p>study finds work requirements study says study shows states expanded health care affordable care low income expansion tied health insurance</p>	<p>deaths every state study could averted every state expanded could averted deaths heart related deaths fewer heart related tied fewer heart expansion tied fewer data show failing show failing expand</p>
c	<p>health insurance food stamps private insurance health care social security make much doesn cover nursing home mental health full time</p>	<p>long term care make much money work full time afford health insurance blue cross blue cross blue shield mental health care pre existing conditions social security disability pre existing condition</p>
i	<p>centers services long term managed care health care term care health insurance nursing home mental health home care behavioral health</p>	<p>long term care critical carecoordination model carecoordination model elements managed care members model elements managed elements managed care protect family assets assets devastating nursing devastating nursing home family assets devastating</p>
n	<p>work requirements health care trump administration judge blocks social security federal judge managed care requirements kentucky security budget kentucky arkansas</p>	<p>work requirements kentucky requirements kentucky arkansas social security budget trump said wouldn wouldn social security said wouldn social security budget cuts judge blocks work judge strikes work voter approved expansion</p>
p	<p>social security food stamps health care health insurance cuts social private insurance</p>	<p>cuts social security pre existing conditions billion social security cutting social security social security cuts welfare food stamps</p>

	trump budget pre existing existing conditions take away	trillion billion billion social security public billion billion social social security budget
--	--	--