Prioritization of Risk Factors for the Prevention and Treatment of Pediatric Obesity

Award Number: 1637312, Start Date: 09/01/2016, End Date: 08/31/2018

Lu Wang[1], Dongxiao Zhu[1] (Co-PI), Elizabeth Towner[2] (Co-PI), Ming Dong[1] (PI)
1: Department of Computer Science, Wayne State University, Detroit, MI 48202
2: Department of Family Medicine & Public Health Sciences, Wayne State University, Detroit, MI 48202

CHALLENGE

Obesity is one of the leading preventable causes of death in the United States (U.S.). Risk factor analysis is a process to identify and understand the risk factors contributing to a particular disease, and is an imperative component in the development of efficient and effective prevention and intervention efforts. Most existing methods usually aim to build a one-size-fits-all model to identify the risk factors at the population-level. However, this type of methods do not take into consideration of population heterogeneity.

SOLUTION

To overcome this limitation, we employ the framework of multi-task learning (MTL) to identify a ranked list of obesity risk factors for each subpopulation/participant (task) simultaneously with utilizing appropriate shared information across tasks. By learning multiple related tasks at the same time, MTL provides a paradigm to rank risk factors at multi-levels.

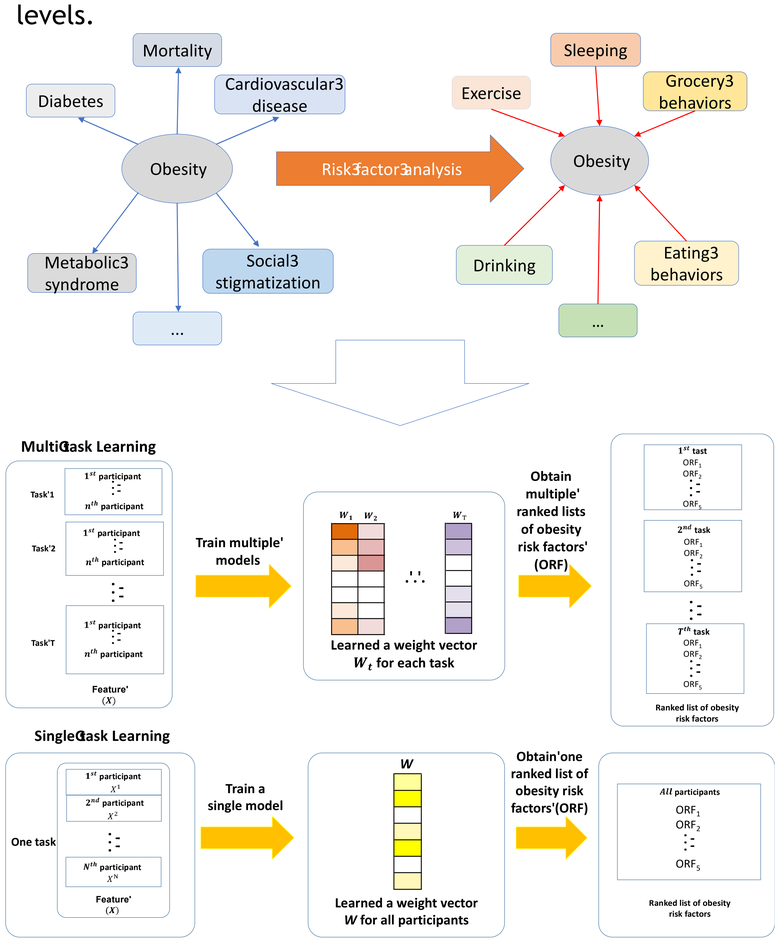

Figure 1. Graphical representation of our idea.

Method

All tasks share a common set of features, $\ell_{2,1}$-norm regularized multi-task learning is able to capture the task relatedness from multiple related tasks:

$$\min_{W} \mathcal{L}(W) + \lambda \|W\|_{2,1}$$

- $W \in R^{G \times T}$: weights.
- T: number of tasks, t: index of task, g: index of feature and G: number of continuous features.
- $\mathcal{L}(W) = \sum_t^T \left\| W_t^T X_t - Y_t \right\|_F^2$, where X_t is the input matrix of tth task and Y_t is the corresponding output variable.
- $\|W\|_{2,1} = \sum_{t=1}^T \|W_t\|_2 = \sum_{t=1}^T \sqrt{\sum_{g=1}^G |w_{gt}|^2}$: the group sparse penalty. λ: tuning parameter.

DATA

	Behavioral risk factor surveillance system (BRFSS)			Pediatric obesity data collected in the Detroit and Metro Detroit (POMM)[1]		
Participants	2,000 adults (Adults aged 18 years and older) living in various states in the U.S.			23 children (Preschoolers aged 2-5 years) living in the Detroit and Metro Detroit		
Data collection	Telephone or land phone interviews			Paper-based surveys		
Input variables	Multi-faced			Multi-levels		
	Demog-raphy	Health	Social life	Child	Family	Community
	Age, geogra-phic inform-ation, family size, etc.	Number of days with bad physical or mental health, years of getting diabetes, etc.	Phone usage, days of worki-ng per week, etc.	Hei-ght, wei-ght, etc.	Lifestyle behavi-ors, grocery behavi-ors, social support, etc.	Neighbor-hood walkability/ safety, transpor-tation access, built environme nt, etc.
Output variable	**Body mass index (BMI)**			**Body mass index z-scores (BMI-z)**		

Table 1. Descriptions of data we use for experiments.

[1]Targeted Research Grant Program, Society of Pediatric Psychology, Division 54 of the American Psychological Association.

RESULTS

Figure 2. Top 5 selected obesity risk factors for each subpopulation using BRFSS data in terms of diverse age groups, i.e., T1 (Young adults:18-44), T2 (Middle-aged adults:45-64), T3 (Older-aged adults: 65-99) and T4 (Age information is missing).

EnergyDaysPerMonth	T1	T2	T3	T4
ExercisePerMonth	T1	T2	T3	
PainNonActiveDaysPerMonth	T2	T3	T4	
WorkHoursPerWeek	T1	T2		
DrinkPerMonth	T1	T4		
HoursSleepPerDay	T2	T3		
AgeOfDiabetes	T1			
AnxiousDaysPerMonth	T3			
PhonesUsage	T4			
PrediabetesBloodSugar	T4			

ExercisePerMonth	AL	AZ	AK	CT	CA	CO	DE	FL	IN	HI	ID	IA	KS	KY	SC	SD	LA	ME	MD	MA	MI	MN	MS
	MO	NH	NE	NV	NJ	NM	NY	UT	TN	TX	WY	WI	WV	WA	RI	PA	OR	OK	OH	ND	NC	VT	VA
DrinkPerMonth	AL	AZ	AK	CT	CA	CO	DE	FL	IN	HI	ID	IA	KS	KY	SC	SD	LA	ME	MD	MA	MI	MN	MS
	MO	NH	NE	NV	NJ	NM	NY	UT	TN	TX	WY	WI	WV	WA	RI	PA	OR	OK	OH	ND	NC	VT	VA
HoursSleepPerDay	AL	AZ	AK	CT	CA	CO	DE	FL	IN	HI	ID	IA	KS	KY	SC	SD	LA	ME	MD	MA	MI	MN	MS
	MO	NH	NE	NV	NJ	NM	NY	UT	TN	TX	WY	WI	WV	WA	RI	PA	OR	OK	OH	ND	NC	VT	VA
PhonesUsage	CA	DE	HI	IA	KY	SD	ME	NE	NV	NJ	NY	TN	WI	RI	VT	VA							
BadPhysicalDaysPerMonth	AK	ID	KY	MA	MS	MO	NH	NE	NV	NJ	TX	OK											
TimesSeeADocPerYear	AZ	FL	SC	MD	NY	UT	TN	TX	WI	RI	VA												
AnxiousDaysPerMonth	AL	IA	SD	MI	WY	PA	OR	OH	NC														
HealthDaysPerMonth	CT	IN	KS	MN	WV	OR	OK																
ChildrenNumber	AZ	IN	LA	MN	WV	ND																	
BadMetalDaysPerMonth	CA	ID	ME	MS	WY																		
DrinkPerDay	CO	KS	SC	MO	WA																		
IncomePerFamily	CO	LA	MI	NM	WA																		
TimesBodyCheckPerYear	AK	MD	NH	UT	PA																		
PainNonActiveDaysPerMonth	FL	MA	NM	ND																			
TimesUgentCarePerYear	HI	OH	VT																				
AsthmaDaysPerYear	AL	NC																					
AdultsNumberPerFamily	DE																						
NoInsuranceMonthPerYear	CT																						

Figure 3.Top 5 selected obesity risk factors for each subpopulation using BRFSS data in terms of various states in U.S.

Age_CG	201	202	204	206	207	209	211	212	213	218	219	220	222	223	224	225	229	230	231	232	234	235	236
total_score	201	202	204	206	207	209	211	212	213	218	219	220	222	223	224	225	229	230	231	232	234	235	236
PSSTotal	201	206	207	209	211	212	213	218	219	220	222	224	225	231	234	235	236						
BIS_Nonplanning	201	202	206	209	211	212	213	218	223	225	230	231	232	234	235	236							
avail_score	202	204	207	213	218	219	222	223	224	225	229	230	231	235									
SNAPamt	206	207	209	212	219	220	222	223	224	230	232	234	236										
BIS_Motor1	211	220	229	232																			
PromispTotal	201	204																					
FFBS_MaternalControl	202	204																					
BIS_Attentional	229																						

Figure 4.Top 5 selected obesity risk factors for each participant using POMM data.

SCIENTIFIC/BROADER IMPACT

- Research impact: transformative to other research areas, e.g., using ranking-CNN in images to estimate the age.
- Educational impact: Data Science program at Wayne State University:
  1. Enhanced student training in a boarder sense via a campus-wide data science program (e.g., at department or college levels);
  2. Multidisciplinary research training, i.e., the students in this and related projects meet with PIs each week to discuss research problems.

CONTACT

Dongxiao Zhu, Ph.D.
Associate Professor
http://www.cs.wayne.edu/dzhu/
dzhu@wayne.edu