ProHealthSu17 - Kate Sanders - Rob Long

[Extended Abstract]

Author One*
Institute One
Address One
author.one@emails.com

Author Two
Institute Two
Address Two
author.two@emails.com

Author Three
Institute Three
Address Three
author.three@emails.com

Author Four
Institute Four
Address Four
author.four@emails.com

ABSTRACT

Abstract text. Abstract text. Abstract text. Abstract text. Abstract text. Abstract text. Abstract text.

Categories and Subject Descriptors

H.4 [**Information Systems Applications**]: Miscellaneous; D.2.8 [**Software Engineering**]: Metrics—*complexity measures, performance measures*

General Terms

Theory

Keywords

ACM proceedings, LaTeX, text tagging

1. INTRODUCTION

A.) Problem

a.) Peoples behaviors have a tendency to drastically change over time.

b.) These short term changes can play a huge role in the patients health status

c.) Short term changes are not easily observable.

d.) If a model were to be created displaying these changes it could help patients make better decisions about their own lives.

B.) Obstacles

a.) creating a complicated longitudinal model is extremely time consuming.

b.) Longitudinal models can sometimes be intractable

c.) There is a lot of background research that needs to be done before creating and implementing such a complex idea

C.) Dynamic Bayesian Networks

a.) We used a Dynamic Bayesian Network because it enabled us to represent the dependence between variables and to give a concise specification of the joint probability distribution. [6]

b.) Within the network each node or variable is equipped with a conditional probability table (CPT). The nodes are then connected by edges that represent direct influence of one node to another.All together, permitting the network to make statistically based inferences, in modeling the CARDIA data set. [6]

c.) Used Bases rule as our guide

i.) $P(A|B) = P(B|A)P(A) / P(B)$

D.) Technology Hole

a.)There has been Dynamic Bayesian Networks used to show connections between behavior and disease b.) No model has been produced that specifically represents behavior

E.) Solution

a.) We have created a preliminary Dynamic Bayesian Network that will aid in tracking and modeling human behavior using CARDIA data.

F.) Contributions

a.) In this paper, we make thing following contributions:

b.) Our research can be used to build a decision support model for hospitals to aid in several tasks including:

*Now on postdoctoral fellow at ABC University i.) resource allocation

ii.) treatment planning

iii.) Prospectively provide physicians with valuable resources required for making informed decisions

c.) Producing a more personalized medical approach for hospitals to use when treating their patients.

d.)The methods we used to model these behavioral changes can serve as a backbone for further modeling related work.

Feature Name	Type	Threshold
Marital Status	Boolean	
Employment Status	Boolean	
Home Ownership	Boolean	
Health Insurance Status	Boolean	
Obesity	Boolean	
Difficulty Paying for Basic Necessities	Boolean	
Sex	Boolean	
Electrocardiogram and Echocardiography	Boolean	
Coronary calcium	Boolean	
Age	Ordinal	18,24,25,30
Race	Ordinal	1,8
Education Level	Ordinal	1,8
Blood Pressure	Ordinal	cell6
Chemistries	Ordinal	
Anthropometry	Ordinal	1,500
Medical history	Ordinal	1,5
Family history	Ordinal	1,5
Physical Activity	Ordinal	1,10
Nutrient intake/dietary history	Ordinal	1,10
Psychosocial parameters	Ordinal	1,10
Pulmonary function	Ordinal	1,3
Carotid intimal medial thickness	Ordinal	1,3
Genetic studies	Ordinal	1,10

2. RELATED WORKS

Previous research with the CARDIA database statistically links behaviors and socioeconomic status with psychosocial vulnerability [8, 7], physical fitness [4, 9], and relative risks of heart disease [3].

Bayesian Networks have been used to model the longitudinal links between behavioral and socioeconomic data and Coronary Artery Calcification [10]. Machine learning techniques have also been used to predict whether an individual has a rare disease based solely on behavioral data [5]. However, this paper contains the first implementation of a Bayesian Network for modeling solely longitudinal behavioral data.

3. METHODS

The data used to create this network was gathered through the Coronary Artery Risk Development in Young Adults (CARDIA) study. The specifics of the study procedures are detailed elsewhere[1]. This study followed 5115 subjects from 1985-6 until present. Birmingham, AL; Chicago, IL; Minneapolis, MN; and Oakland, CA served as centers for data collection. Each location recruited participants evenly distributed between the subgroups of gender, race, education level, and age-group (18-25 or 25-30). Data gathered from participants included physical measurements, clinical tests, and an in depth questionnaire about lifestyle and socioeconomic status. This study has served as the backbone for major research in the development of heart disease.

We consider only the longitudinal lifestyle and socioeconomic data in our research. Socioeconomic factors included age, sex, race, education level, marital status, employment status, difficulty paying for basic necessities, home ownership, and health insurance status.

To model how behaviors change over time, we applied the CARDIA data to BN learning algorithms, which used a hill-climbing approach. To score structures we use BIC, Mutual information, and BDe, [2] because these metrics have shown promising in previous BN research [10].

4. RESULTS AND DISCUSSIONS

5. REFERENCES

[1] G. D. Friedman, G. R. Cutter, R. P. Donahue, G. H. Hughes, S. B. Hulley, D. R. Jacobs, K. Liu, and P. J. Savage. Cardia: study design, recruitment, and some characteristics of the examined subjects. *Journal of Clinical Epidemiology*, 41(11):1105–1116, jan 1988.

[2] D. Heckerman, D. Gelger, and D. M. Chickeriag. Learning Bayesian Networks: The Combination of Knowledge and Statistical Data.

[3] A. S. Karlamangla, B. H. Singer, D. R. Williams, J. E. Schwartz, K. A. Matthews, C. I. Kiefe, and T. E. Seeman. Impact of socioeconomic status on longitudinal accumulation of cardiovascular risk in young adults: the CARDIA Study (USA). *Social Science & Medicine*, 60(5):999–1015, 2005.

[4] C. E. Lewis, D. E. Smith, D. D. Wallace, O. D. Williams, D. E. Bild, D. R. Jacobs, and Jr. Seven-year trends in body weight and associations with lifestyle and behavioral characteristics in black and white young adults: the CARDIA study. *American journal of public health*, 87(4):635–42, apr 1997.

[5] H. MacLeod, S. Yang, K. Oakes, K. Connelly, and S. Natarajan. Identifying Rare Diseases from Behavioural Data: A Machine Learning Approach. In *Proceedings - 2016 IEEE 1st International Conference on Connected Health: Applications, Systems and Engineering Technologies, CHASE 2016*, 2016.

[6] S. J. Russell and P. Norvig. *Artificial Intelligence: A Modern Approach*, volume 9. 1995.

[7] L. W. Scherwitz, L. L. Perkins, M. A. Chesrtey, G. H. Hughes, S. Sidney, and T. A. Manolio. Hostility and Health Behaviors in Young Adults: The CARDIA Study. *American Journal of Epidemiology*, 136(2):136–145, jul 1992.

[8] G. Scherwitz, L; Perkins, L; Chesney, M; Hughes. Cook-Medley Hostility scale and subsets: relationship to dem... : Psychosomatic Medicine. *Psychosomatic Medicine*, 53(1):36–49, 1991.

[9] M. H. Shishehbor, P. Gordon-Larsen, C. I. Kiefe, and D. Litaker. Association of Neighborhood Socioeconomic Status with Physical Fitness in Healthy

Young Adults: the CARDIA Study.
[10] S. Yang, K. Kersting, G. Terry, J. Carr, and S. Natarajan. Modeling Coronary Artery Calcification Levels From Behavioral Data in a Clinical Study.