IJCAT International Journal of Computing and Technology, Volume 1, Issue 7, August 2014
ISSN : 2348 - 6090
www.IJCAT.org

A Framework for Generating the User's Behavioral Consequences for Predicting Outcomes

[1] K. Vasantha , [2] K. Narayana

[1] Post-Graduate Student, Department of Computer Science and Engineering
SIT, PUTTUR, India

[2] Head & Associate Professor, Department of computer Science and Engineering
SIT, PUTTUR, India

Abstract - To retrieve the data from plenty of information and finding exactly what data to get mined has become progressively computerized. Alternatively selecting what data to get together requires human association or practice, generally delivered by field expert. This technique describe that for prediction of some behavioral outcomes, non-field experts can jointly formulate structures after which provide values. This paper gives new strategy to machine science demonstrating that non domain experts can collectively formulate features and provide values for all those features so they really are predictive of some behavioural outcome interest. This is accomplishing by web platform where crowd get connected to one another by answering and adjusting question that assist to predict behavioural outcome. Which result in *dynamically* growing online survey.

Keywords **- Crowd sourcing, Human Behavior modeling, Survey, Prediction**.

1. Introduction

These days there are various problems through which exact option would be extremely hard. In these cases you have to calculate the consecutive result. Such issues a team of experts becomes necessary for each and every individual domain which ends excessively decrease in human efforts. As an example, the survey designer need to be an expert of these domain to select appropriate questions related respected domain. An engineer must keep correlation and recognized approach of design So that you can judge which concept may well be more efficient such a way that it's going to improve the performance. Necessity of domain expert would be the main problem with this approach. However, when using the information about crowd to understand the difficult problems will harness the potency of result. Thus, with regards to this web-based approach is usually to achieve active participation of crowd in suggesting questions together with providing solutions to the given questions that leads to continuing development of predictive model.

To produce predictive models between predictor variables and an outcome, there are numerous problems. Once the number of predictive covariates as well as the model structure are pre-specified, tool like neural networks provide advanced techniques for calculating model parameters. Now, current principals are providing new tools for gathering the primary kind of non-linear predictive models of given good input -output data. Though, the position of selecting which potentially predictive variables to analyze is essentially a qualitative task that will require essential domain expertise. As an example, to decide on questions that could find predictive covariates a survey designer has to have domain expertise. In order to find which variables might be logically adjusted also to optimize performance an engineer must improve significant understanding of a design.

2. Related Work

In order to develop predictive model which map between group of outcomes and predictor variables create many problems. Even though the model structure are pre-specified with group of predictive covariates statistical tools provides mature techniques to compute model parameter. The objective of the analysis should be to test a different approach to modelling through which online crowd enables you to define potentially predictive variables to review by asking and receiving respond to question, to ensure that a predictive model is developed.

2.1 Machine Science

Machine science is usually a new scientific technology which worth it to read to uncover, analyses, classify The

information to create hypothesis and develop models Machine Science is usually a hot topic Within the theory and philosophy of recent science, with recent claims that "with a decade, all the more powerful tools will enable automated, high-volume hypothesis generation to help high-throughput experiments in biomedicine, chemistry, physics, and in some cases the social sciences".

This paper introduces a technique during which non domain experts are usually motivated to formulate independent variables and also populate an adequate amount of variables to create successful modelling. This could be well explained the following. Users or user go to the site dependant on behavioural outcomes (The behavioural outcomes can be quite a body mass index or daily electricity consumption) is usually to be modelled.

The user will give you their particular outcomes (similar to their own utilization of electricity) and answer the questions which can be predictive of their outcome (Including simply how much electricity they choose daily).By ordinarily, the latest models of are constructed in oppose to growing data sets predicting user's behavioural outcome. User may also post their particular questions that, which becomes new independent variables when answer by other users within the modeling process. Thus to find out and populate independent variables are going to be made by user community.

2.2 Crowd Sourcing

The rapid development in user-generated content on-line is usually an example of how bottom-up interactions can, under some circumstances, effectively solve issues that previously required explicit management by teams of experts. Harnessing the ability and of enormous variety of individuals can often be generally known as "crowd sourcing" and contains been used effectively in numerous research and commercial applications.

Although arguably not strictly a crowd sourced system, the rapid rise of Wikipedia illustrates how online collaboration may be used to solve difficult problems (the advance of an encyclopedia) without financial incentives. several reviews crowd sourcing tools and argues that direct motivation tasks (tasks through which users are motivated to execute the position since they realize its useful, as an alternative to for financial motivation) can produce results which have been finer quality than financially motivated tasks. Similarly, reports that levels of competition are useful in improving performance using a task with either direct or indirect motivation.

3. Methodology

The position of choosing which potentially predictive variables to examine is essentially a qualitative task that will require substantial domain expertise. For example, a survey designer should have domain expertise to select questions that may identify predictive covariates. An engineer must develop substantial understanding of a design so that you can determine which variables might be systematically adjusted so that you can optimize performance.

It described here wraps an individual behavior modeling paradigm in cyber infrastructure in a way that:

(1) The investigator defines some human behavior-based outcome that may be for being modeled

(2) Information is collected from human volunteers

(3) Models are always generated automatically

(4) The volunteers are motivated to propose new independent variables.

Architecture:

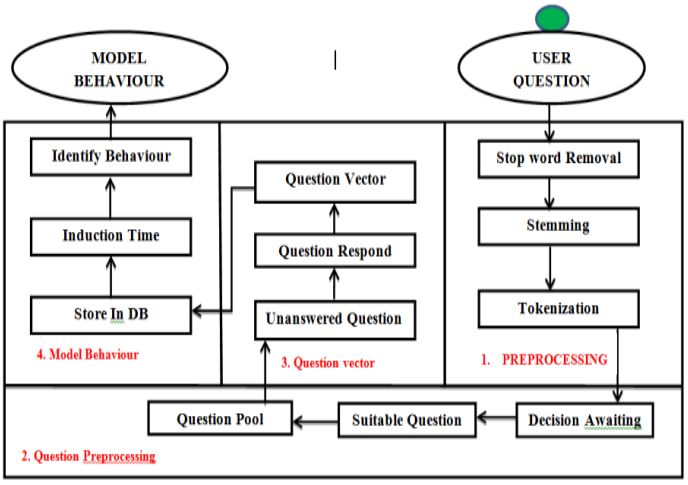

Fig 1

The proposed system shown in above Fig. Where the technique is divided into four parts which might be

1) Pre-processing,

2) Question Pre-processing,

3) Question vector and

4) Model Behavior.

User is key to it, the system starts off with user question, through which user may easily respond the question to calculate a behavioral outcomes and find the suitable result. Non-domain experts may responds to questions.
Making sure that first input or data set {for our for the} technique is that user natural thinking by means of natural/general questions. These questions are further handle from the other places of system which described in more detail below.

1) Pre-Processing:

Pre-processing could be the initial step in a very system that can take The user input such as natural language and evaluate that input in a suitable format have for the next phase of system that's question Pre-processing. Pre-processing perform the essential operation on users input question.

Following steps mixed up in Pre-processing

a) Stop word removal
b) Stemming
c) Tokenization

2) Question Pre-Processing

The question revealed the identity of that author (e.g. “Hi, I am John Doe. I'd really like to recognize if...”) thereby contravening the Institutional Review Board approval for these experiments; The question contained profanity or hateful text; The question was inappropriately correlated With all the outcome (e.g. “Precisely what is your BMI?”). Should the question was deemed suitable it had been included in the pool of questions entirely on the site; otherwise the question was discarded.

3) Question Vector

Users respond the question which might be already loaded in site and as well pose their new question to others that unanswered questions is used question vector additional processing.

4) Model Behavior

Users who visit the site first provide their individual value for that results of interest. Users will then respond to questions located on the site. Their answers are kept in a standard data set making it accessible to the modeling engine.

4. Conclusion

This paper provides A brand new strategy to social science modeling through which human behavioral outcome is generated by motivating the participants. In this particular paper participate is coming to the web site and answering towards the questions which he really wants to instead of wishes to be brings about hectic with the participant. So through the use of principle approach while using the different regression model question ordering may very well be made principally in order that the user couldn't face the questions that they don’t would like to answer.

Also rather then while using the single model the significant with the system could possibly be enhance by making use of the n variety of models. This can be the potentially an alternative way to accomplish science. This new strategy to science could result the exponential growth which can be found tin another online collaborative communities.

References

[1] Josh C. Bongard, “Crowdsourcing Predictors of Behavioral Outcomes,” IEEE transactions on knowledge and data engineering, 2013.

[2] S. Bowman, S. Gortmaker, C. Ebbeling, M. Pereira, and D. Ludwig, “Effects of fast-food consumption on energy intake and diet quality among children in a national household survey,” Pediatrics, vol. 113, no. 1, p. 112, 2004.

[3] J. Currie, S. Della Vigna, E. Moretti, and V. Pathania, “The effect of fast food restaurants on obesity and weight gain,” American Economic Journal: Economic Policy, vol. 2, no. 3, pp. 32–63, 2010.

[4] D. C. Brabham, “Crowdsourcing as a model for problem solving,” Convergence, vol. 14, pp. 75–90, 2008.

[5] Paul Whitla, “Crowdsourcing and Its Application in Marketing Activities”, Contemporary Management Research Pages 15-28, Vol. 5, No. 1, March 2009.

[6] J. Giles, “Internet encyclopedias go head to head,” Nature, vol. 438,no. 15, pp. 900–901, 2005.

[7] D. C. Brabham, “Crowdsourcing as a model for problem solving,”Convergence, vol. 14, pp. 75–90, 2008.

[8] A. Sorokin and D. Forsyth, “Utility data annotation with amazon mechanical turk,” in Proc. IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops, 2008.

[9] M. Marge, S. Banerjee, and A. Rudnicky, “Using the amazon mechanical turk for transcription of spoken language,” in Proc. IEEE International Conference on Acoustics Speech and Signal Processing,2010.

[10] N. Kong, J. Heer, and M. Agrawala, "Perceptual guidelines for creating rectangular treemaps," IEEE Transactions on Visualization and Com- puter Graphics, vol. 16, no. 6, 2010.

[11] A. Kittur, E. Chi, and B. Suh, "Crowdsourcing user studies with mechanical turk," in Proc.Twenty-sixth annual SIGCHI conference on human factors in computing systems, 2008.

[12] D. Wightman, "Crowdsourcing human-based computation," in Proceed- ings of the 6th Nordic Conference on Human-Computer Interaction: Extending Boundaries, 2010.

[13] B. Fitzgerald, "The transformation of open source software," Manage- ment Information Systems Quarterly, vol. 30, no. 3, pp. 587–598, 2006.

[14] J. Howe, Crowdsourcing: Why the Power of the Crowd is Driving the Future of Business. Crown Business, 2009.

[15] N. Thurman, "Forums for citizen journalists? adoption of user generated content initiatives by online news media," New Media and Society, vol. 10, no. 1, 2008.

[16] C. DiBona, M. Stone, and D. Cooper, Open Source 2.0: The Continuing Evolution. O'Reilly Media, 2005.

[17] J. Leskovec, L. Adamic, and B. Huberman, The Dynamics of Viral Marketing. New York: ACM Press, 2007.

[18] K. Lerman, "Social networks and social information filtering on digg," arXiv: cs/0612046v1, 2006.

[19] D. Anderson, J. Cobb, E. Korpela, M. Lebofsky, and D. Werthimer, "Seti@home: an experiment in public-resource computing," Communi- cations of the ACM, vol. 45, no. 11, pp. 56–61, 2002.

[20] J. Cohn, "Citizen science: Can volunteers do real research?" BioScience, vol. 58, no. 3, pp. 192–197, 2008.

[21] J. Silvertown, "A new dawn for citizen science," Trends in Ecology & Evolution, vol. 24, no. 9, pp. 467–471, 2009.

[22] A. Beberg, D. Ensign, G. Jayachandran, S. Khaliq, and V. Pande, "Folding@home: Lessons from eight years of volunteer distributed computing," in IEEE International Symposium on ParallelDistributed Processing, May 2009, pp. 1–8.

[23] C. Lintott, K. Schawinski, A. Slosar, K. Land, S. Bamford, D.Thomas, M. Raddick, R. Nichol, A. Szalay, D. Andreescu et al., "Galaxy zoo: morphologies derived from visual inspection of galaxies from the sloan digital sky survey?" Monthly Notices of the Royal Astronomical Society, vol. 389, no. 3, pp. 1179–1189, 2008.

[24] S. Cooper, F. Khatib, A. Treuille, J. Barbero, J. Lee, M. Beenen, A.Leaver-Fay, D. Baker, Z. Popovi´ c et al., "Predicting protein structures with a multiplayer online game," Nature, vol. 466, no.7307, pp. 756– 760, 2010.

[25] A. Kittur, "Crowdsourcing, collaboration and creativity," XRDS, vol. 17, no. 2, pp. 22–26, 2010.