ISSN: 2321-7782 (Online)
Volume 2, Issue 11, November 2014
International Journal of Advance Research in Computer Science and Management Studies
Research Article / Survey Paper / Case Study
Available online at: www.ijarcsms.com

A Survey on Forecasting Behavioral Outcome through Crowdsourcing Mechanism

Mane Kush Chandrakant[1]
Department of Computer Engineering,
DGOI, FOE, Daund.
Pune, Maharashtra – India

Bere Sachin Sukhadeo[2]
Department of Computer Engineering,
DGOI, FOE, Daund.
Pune, Maharashtra – India

Abstract: ***Producing models from large data sets and determining which subsets of belonging to data is becoming increasingly automated. However selecting what fact to collect in the first place requires human intuit or experience, usually supplied by a domain skill. we describe a new approach to machine science which prove for the first that non-domain experts can collectively formulate features, and provide values for those features such that they are prophecy of some to conduct outcome of interest. This was accomplished by building a web platform in which human groups interact to both respond to questions likely to help forecast to conduct result and sit new questions to their peers. This result in a dynamically growing online observe, but the result of this cooperative to conduct also leads to models that can forecasts users result based on their responses to the user generated observer query. Here we describe two web based experiments that instantiate this approach, the one is monthly electric energy consumption and other body mass index.***

Keywords: Crowdsourcing, machine science, surveys, social media. Human behaviour modelling

I. INTRODUCTION

There are numerous problem in which one solutions to create prophecy models to map between a set of forecasting variables and an result. Statistical tools such as multiple regression or neural networks provide to reach methods for calculating model parameters when the set of prophecy covariates and the model structure are pre-defined. Furthermore current research is providing new tools for inference the structural from of non-linear forecasting models, given good input and output data. However the task of alternating which potentially forecasting variables to study morally a worth task that requires substantial ruled over expertise. For example a designer look must have domain expertise to alternate question that will identify forecasting covariates. An engineer must mature substantial familiarity with a design in order to calculate which variables can be systematically adjusted in order to finding performance.

The need for including of domain experts can become a bottleneck to new insights. However if the teeth of crowds could be harnessed to create insight into difficult problems, one might see exponential rises in the discovery of the casual factors of behavioral result mirroring the exponential growth. Thus the aim of this research was to test an alternative approach to modelling in which the wisdom of crowds is harnessed to both propose potentially forecaster variables to study by asking query, and respond to those query, in order to produce a predictive model.

II. RELATED WORK

1) Machine science

The MS is an increasing trend that conducts to automate as numerous aspects of the scientific method as possible. From long data history, Automated production of models that are currently robot scientists have been proved that can physically carry out test as well as hypothesis production, experimental design, experiment execution, and hypothesis refutation. In a forecasting problem, MS is not yet able to select the nondependent variables that might forecast an result of interest, and for which data

collection is required. However one aspect of the scientific method that has not yet yielded to automation is the selection of variables for which data should be collected to calculate supposition.

This paper introduces, for the first time, a method by which non domain skill can be motivated to formulate independent variables as well as populate enough of these variables for successful modelling. In short, this is accomplished as follows. Users reach at a website in which to conduct result is to be modelled. First of all Users provide their own outcome and then answer query that may be forecasting of that result. Periodically, models are built against the growing data set that predict each user's to conduct the result. Users may also sit their own questions that, when answered by other users, become new independent variables in the modelling process.

2) Crowdsourcing

The fast growing user produced data on Internet is an example of crowdsourcing this is very helpful where previously a group of experts is nesessery. Harnessing the experience and hardworking of large numbers of individuals is known as "crowdsourcing".

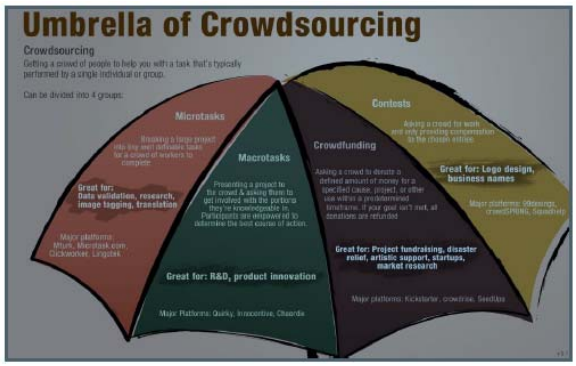

Fig.1 Umbrella of Crowdsourcing

As shown in above diagram gives us the idea about how crowdsourcing works and how it is effective in many fields now days. In many cases when we have some problems, each time we have to go to domain experts for solution of our problem. This problem can be solve with the help of crowdsourcing i.e. user can solve their own problems by their own selves. The best example which proves the effectiveness of crowdsourcing is Amazon's Mechanical Turk. In this one can explain a "Human Intelligence Task" such as characterizing data, transcribing spoken language, or creating data visualizations with the help of group of people which is very difficult for a computer alone.

In a crowd sourced system, the Wikipedia illustrates how online collaboration can be used to solve difficult problems without financial incentives. This are two tasks with direct motivation: for the body mass index task, users are motivated to understand their lifestyle alternate in order to approach a healthy body weight for the household energy usage task, users are motivated to understand their home energy usage that means to increase their energy efficiency. they compare with other participants users and by ranking the prophecy quality of questions that participants user provide. In the literature and commercial applications that laypersons are more willing to respond to reach and queries from peers than from authority figures or organizations.

In comparison with the top-down system the collaborative systems are generally more scalable. Crowdsourcing can tend to develop a creative solution that is substantially different from the experts. The crowd sourced poem translation task is surprising and preferable than the expert translation.

III. METHODOLOGY

This system described here wraps a human behaviour modelling paradigm in cyber infrastructure such that

(1) The investigator defines some people to conduct the result that is to be modelled;

(2) Data is collected from human done willingly;

(3) Small copies are continually produced automatically; and

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

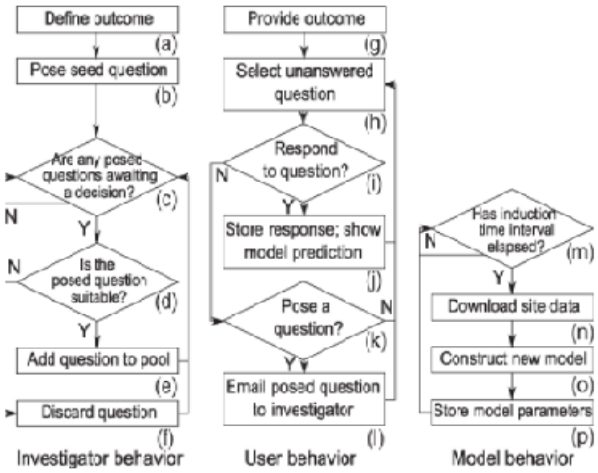

Figure 2. Overview of the system.

As shown in above Fig. The investigator (**a-f**) is responsible for creating the web platform initially, and seeding it with a starting question. The user generated by new query as they filter runs experiments. Users (**g-l**) may elect to answer as-yet unanswered observe query or sit some of their own. The small part of engine (**m-p**) continually produce prophecy copies using the observe query as candidate forecasting of the result and users' responses as the training data.

1) Investigator Behavior

It is responsible for initially creating the web platform, and seeding it with a starting query. Then, as the observation runs they filter new survey query produced by the users.

However, once sit, the query was abled by the investigator as to its suitability. A query was deemed unsuitable if any of the following conditions were met:

1. The query revealed the identity of its author (e.g. "*Hi, I am kush. I could look to know if...*") thereby contravening the Institutional Review Board approval for these test;
2. The query contained hateful text;
3. The query was inappropriately correlated with the result.

If the query was deemed suitable it was added to the pool of query available on the site; otherwise the query was removed.

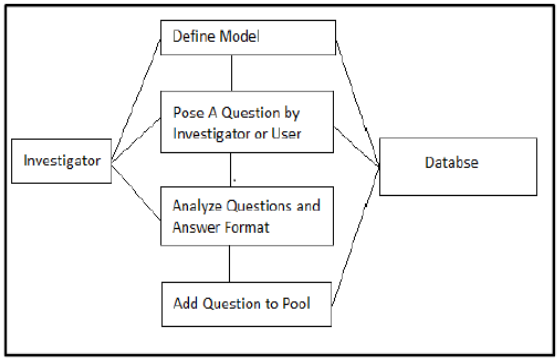

Fig. 2 Investigator Model

2) User Behavior

Users who visit the site first provide their individual value for the result of interest. Users may then respond to queries found on the site. Their answers are stored in a common database and made available to the modeling engine.

At any time a user may elect to sit a query of their own devising. Users could sit queries that required a yes/no response, a five-level Like rating, or a number. Users were not constrained in what types of query to sit.

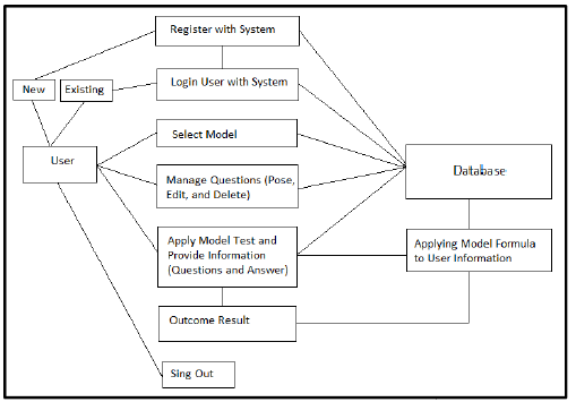

Fig. 3 User Model

3) Model Behavior

The modeling engine continually produce prophecy models using the observer queries as candidate forecaster of the result and users' responses as the training data.

IV. CONCLUSION

In This paper we have studied a new approach to social science modeling in which the participants themselves are motivated to uncover the correlates of some human to conduct result, such as homeowner electricity usage or body mass index. In both cases participants successfully uncovered at least one statistically significant forecaster of the result variable. For the body mass index outcome, the participants successfully formulated many of the correlates known to prophecy BMI, and provided sufficiently honest values for those correlates to become forecasting during the test. While, our instantiations focus on energy and BMI, the proposed method is general, and might, as the method improves, be useful to answer many difficult questions regarding why some result are different than others. For example, future instantiations might provide new insight into difficult query like: "Why do grade point averages or test scores differ so greatly among students?", "Why do certain drugs work with some populations, but not others?", "Why do some people with similar skills and experience, and doing similar work, earn more than others?"

Despite this initial success, much work remains to be done to improve the functioning of the system, and to validate its performance.

ACKNOWLEDGMENT

I express great many thanks to Prof. Sachin S. Bere for their great effort of supervising and leading me, to accomplish this fine work. To college and department staff, they were a great source of support and encouragement. To my friends and family, for their warm, kind encourages and loves. To every person gave us something too light my pathway, I thanks for believing in me.

References

1. J. Bongard and H. Lipson, "Automated reverse engineering of nonlinear dynamical systems," Proceedings of the National Academy of Sciences, vol. 104, no. 24, pp. 9943–9948, 2007.
2. J. Evans and A. Rzhetsky, "Machine science," Science, vol. 329, no. 5990, p. 399, 2010.
3. R. D. King, K. E. Whelan, F. M. Jones, P. G. K. Reiser, C. H. Bryant, S. H. Muggleton, D. B. Kell, and S. G. Oliver, "Functional genomic hypothesis generation and experimentation by a robot scientist," Nature, vol. 427, pp. 247–252, 2004.
4. R. King, J. Rowland, S. Oliver, M. Young, W. Aubrey, E. Byrne, M. Liakata, M. Markham, P. Pir, L. Soldatova et al., "The automation of science," Science, vol. 324, no. 5923, p. 85, 2009.
5. J. Bongard, V. Zykov, and H. Lipson, "Resilient machines through continuous self-modeling," Science, vol. 314, pp. 1118–1121, 2006. [6] J. Giles, "Internet encyclopedias go head to head," Nature, vol. 438, no. 15, pp. 900–901, 2005.
6. D. C. Brabham, "Crowdsourcing as a model for problem solving," Convergence, vol. 14, pp. 75–90, 2008.
7. 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.
8. 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.
9. N. Kong, J. Heer, and M. Agrawala, "Perceptual guidelines for creating rectangular treemaps," IEEE Transactions on Visualization and Computer Graphics, vol. 16, no. 6, 2010.

AUTHOR(S) PROFILE

Mr. Mane Kush C. Received his B.E. degree in Information Technology (Distinction) in the year 2011 from solapur University. He is currently working toward the M.E. Degree in Computer Engineering from the University of Pune, Pune. He has 02 years of teaching experience at undergraduate level. His research interests lies in Data Mining.

Prof. Bere Sachin S. received his B.E. degree in Technology (First-class) in the year 2008 from Pune University and M.Tech Degree (Distinction) in Computer Engineering in2013 from JNTU University. He has 07 years of teaching experience at undergraduate and postgraduate level. Currently he is working as Assistant Professor and HOD in Department of Computer Engineering of DGOI, FOE, swami-chincholi, Daund, Pune University. His research paper has been published in IJTITCC, IJISET year 2014. His research interests are Digital Image processing.

