
Article:
Andrews, J.A., Harrison, R.F., Brown, L.J.E et al. (8 more authors) (2017) Using the NANA toolkit at home to predict older adults' future depression. Journal of Affective Disorders. ISSN 1573-2517

https://doi.org/10.1016/j.jad.2017.02.019

Using the NANA toolkit at home to predict older adults' future depression

Andrews JA[1], Harrison RF[1], Brown LJE[2], MacLean LM[3], Hwang F[4], Smith T[5], Williams EA[1], Timon C[6], Adlam T[7], Khadra H[8], & Astell AJ[1,9]

[1]University of Sheffield, UK

[2]University of Manchester, UK

[3]University of Glasgow, UK

[4]University of Reading, UK

[5]Generic Robotics Ltd, Reading, UK

[6]University College Dublin, Ireland

[7]Designability, Bath, UK

[8]Oxford Technologies Ltd, UK

[9]Ontario Shores Centre for Mental Health Sciences, Canada

Corresponding Author: Jacob Andrews, Jacob.andrews@sheffield.ac.uk, School of Health and Related Research, Regents court, 30 Regent Street, Sheffield, S1 4DA

This article is published online at the following doi: http://dx.doi.org/10.1016/j.jad.2017.02.019

Using the NANA toolkit at home to predict older adults' future depression

Introduction

Depression affects approximately 350 million people worldwide and is "the leading cause of disability" according to the World Health Organisation (WHO 2016). Despite the availability of effective treatments, depression is undertreated owing to a range of barriers including misdiagnosis and social stigma making people reticent to seek help (Préville et al., 2015). This is especially so among older adults, where depression is more common than dementia (Allan et al., 2014). Medical practitioners may find depression harder to identify in this population, particularly when older adults present with more somatic symptoms or report subjective change in functioning. However, multiple, age-related conditions including stroke and diabetes, as well as social factors such as bereavement or social isolation, increase depression risk in later life (Allan et al., 2014).

Traditional measures of depression rely on retrospective recollection of mood and feelings over a one or two-week period. This method is susceptible to bias or distortions based on current mood or memory problems (Brown & Astell 2012). The increasing availability of digital technologies to support data collection has inspired new approaches that can be used outside of the laboratory or clinic. For example, 'momentary' measures that prompt participants to record aspects of their current experience, such as mood or social exchanges, at different times of day (Cain et al., 2009). These can help to avoid the problems associated with retrospective mood measures by capturing the current mood state (Brown & Astell 2012). Technology also provides the potential to collect information from older adults at home over a prolonged period of time, which could identify meaningful patterns or changes in health status or behavior (Kaye et al. 2011).

We developed the Novel Assessment of Nutrition and Ageing (NANA) (Astell et al. 2014) assessment

tool to capture reliable and valid data from older adults in their own homes without a researcher or clinician being present. NANA includes new measures of mood (Brown et al. 2016) and cognitive function (Brown et al. 2016), developed for recording every day, alongside dietary information (Astell et al. 2014). In this paper we explore the use of a machine learning technique with mood data collected via the NANA touchscreen assessment system in older adults' own homes to produce an algorithm for early detection of depression symptoms in older adults. The intended use of the algorithm is as a tool for prediction, imagined to be used with older adults known to be at risk of depression, for example those who have had a stroke. Applications for monitoring may also possible.

Methods

Study Design

This was a retrospective study using data initially collected for the validation of the NANA system (Astell et al. 2014). The data we used included mood scores self-reported by 40 older adults (aged 65+), comprising ratings each day of how 'happy', 'sad', 'tired', 'alert', 'relaxed' and 'hungry' each participant was on a scale of 0-10. Participants completed this self-report daily over three individual weeks (Period 1, Period 2, Period 3), with approximately two-weeks rest between each data collection period. As part of the NANA validation protocol (Astell et al. 2014), participants also completed the Geriatric Depression Scale (GDS; Yesavage & Sheikh 1986; Yesavage et al. 1982) before Period 1 (GDS1) and two weeks after Period 3 ended (GDS2). The GDS2 scores were coded as one or zero according to whether they were above (one) or below (zero) a threshold score of five on the GDS, which is suggested as an indication of depression by Yesavege and Sheikh (1986). This research project received a favourable ethical opinion from the Fife and Forth Valley Research Ethics Committee (08/50501/104). All participants gave written informed consent.

Data

Input variables for this analysis were mean averages of six mood scores (alertness, happiness, tiredness, sadness, relaxation, hunger), over Period 1 (seven days) for each of the 40 participants. The outcome variable was depression status according to GDS scores at the end of the trial, ten weeks after the collection of these first mood reports. The reason for using mood scores from the first period of data collection was to produce a model with the earliest possible predictive ability of later depression status. Three participants' data were excluded during data cleaning due to missing mood data or missing GDS scores. Data points for each of the 37 participants comprised an average mood score for each of the six mood words, plus a depression status (1 or 0). The data were analysed by the first and second authors.

Analysis

We used a least absolute shrinkage and selection operator (LASSO) (Tibshirani, 1996) to explore the possibility of predicting depression status from self-reported mood data from older adults. LASSO is an operation that introduces bias into the estimated coefficients of a (here, logistic) regression procedure by shrinking their absolute size (Tibshirani, 1996). This procedure provides knowledge of which variables are most informative and omits coefficients for those variables which contribute little to the model, resulting in improved predictive performance.

To find the best performing model a "shrinkage" parameter, lambda, must be set, and this was chosen using a repeated stratified cross-validation framework – the currently accepted best practice in machine learning. Cross-validation permits the use of all data in estimation and testing of models and guards against the tendency of more complex models to overspecialise. This is especially important when there are a relatively large number of variables and a small sample size as is the case here. In *k*-fold cross-validation, the data are divided into *k* non-overlapping sets. At each stage, one set is held-out for comparison, while a model is fitted to the remainder. This is repeated for each of the *k* holdout sets. In this study we used cross-validation for parameter selection, setting k=5

(larger values would not allow each hold-out fold to contain a positive outcome). In addition, two hundred Monte Carlo repetitions were made of the cross validation procedure, to guard against the possibility of a favourable random 5-fold partition. We used 100 different values of lambda within the procedure to determine the level at which the usual goodness-of-fit measure, deviance, was minimized, then chose the greatest value of lambda that lay within one standard error of this point (Breiman et al, 1984; Hastie et al, 2009; Krstajic et al, 2014). We justify this on the basis of parsimony: here the increase in deviance associated with the choice is negligible.

Predictive ability was assessed by examining a receiver operating characteristic (ROC) curve generated from predictors of GDS status (according to the model based on 37 records of mood variables), compared against actual GDS scores. All results provided below are cross-validated values owing to the small sample size and the consequent infeasibility of dividing the data to provide an independent out-of-sample set.

Results

Using a LASSO with logistic regression procedure, the deviance was minimized with a lambda value of 0.062 (Figure 1). The largest value of lambda where the deviance fell within one standard error of the minimum was 0.075. At this value of lambda, the model included three variables, ('sadness', 'tiredness' and 'alertness'; see figure 2) as well as a constant. However, the variable contributing least ('alertness') has a relatively small (normalised) coefficient value and, we believe that the practical advantage in reducing the demands on the client group outweighs the cost of a further increase of deviance. These choices resulted in less than a 1% overall increase in deviance. The final model retains only the coefficients for 'sadness' and 'tiredness' corresponding to the largest value of lambda (0.082) that excludes 'alertness'.

Insert Figure 1 about here

The resulting model contained a constant (-2.99) and coefficients for two of the six mood variables, these being tiredness (0.34) and sadness (0.09). Coefficients for sadness and tiredness were both positive, suggesting that higher average sadness and tiredness scores over Period 1 were associated with higher risk of scoring positively for depression at GDS2 in this sample.

Insert Figure 2 about here

The results of the ROC analysis can be found in figure 3. The AUC for the LASSO with logistic regression was 0.88, with a confidence interval of 0.69 to 0.97.

Insert Figure 3 about here

Discussion

We used the LASSO method with logistic regression to produce a predictive indicator of depression status based on data from touchscreen mood self-reports of older adults. The model produced an area under the ROC curve of 0.88, indicating good predictive ability for this sample. While this area is a good portmanteau measure of performance, a particular clinical setting would require a choice of operating threshold at which to make a decision. The ROC curve can be used to determine, for example, the acceptability or otherwise of the sensitivity (true positive rate) associated with a chosen false alarm rate (false positive rate or one-minus-specificity).

The LASSO procedure is designed to reduce the number of variables necessary to produce a model, and to optimize the size of the coefficients included in the model. Of the six mood score variables used here (alertness, happiness, hunger, sadness, tiredness, relaxation), two were retained in the final model (sadness and tiredness). This result suggests that a higher average score for tiredness

and a higher average score for sadness combined were indicative of later positive depression status. The Diagnostic and Statistical Manual of Mental Disorders (American Psychiatric Association 2013) states that low mood and lack of energy are symptoms of depression. Our finding would suggest that these symptoms can be detected prospectively through the use of a daily self-report in this sample, suggesting potential for prediction of later onset of depression.

Our study has a number of strengths. Using a LASSO procedure optimizes the model by including only a limited number of variables. This creates parsimony and reduces the burden of data collection on users. Use of cross-validation creates a 'closed loop' in the development of the model, which focuses attention on predictive ability rather than inference. The AUC of 0.88 (C.I. 0.69-0.97) is high in predictive terms. By applying a machine learning approach, we have created a model that has the potential to predict depression status in older adults through the use of a simple interface.

Limitations

There are limitations to the study. The model produced in this study has been trained on just 37 individuals from two English cities. In its current form it is therefore unclear how well it would apply in clinical use, though the aim of this study was not to produce a clinical tool. Rather, we have demonstrated a technique which could be used with a larger and more representative sample to produce a model which could in turn be put into clinical use with people at risk of a mood disorder.

Early detection of such a condition could allow early intervention in line with national or international guidelines, for example those offered by the World Health Organisation in their Mental Health Gap Action Programme intervention guide (WHO, 2010), potentially reducing the number of older adults who develop a mood disorder and/or commit suicide.

There are also limitations to the analysis we have used. The model is tested on the original sample so results are open to criticism. However, the use of repeated 5-fold cross-validation serves to ameliorate the problem of over-fitting and variance stabilisation and is known to offer conservative

results (Rodriguez et al. 2010) suggesting that, in prospective use, performance might exceed that quoted.

Conclusion

This is the first time to our knowledge that the logistic model, in conjunction with the LASSO has been used to produce an algorithm for prediction of depression status based on older adults' self-reporting of their mood using a touchscreen interface.

Acknowledgements

This work was supported by grant number ES/G008779/1 awarded to Arlene Astell, Elizabeth Williams, Faustina Hwang and Tim Adlam from the New Dynamics of Ageing (NDA) Programme. The NDA funding partners are: Economic and Social Research Council (ESRC), Engineering and Physical Sciences Research Council (EPSRC), Biotechnology and Biological Sciences Research Council (BBSRC), Medical Research Council (MRC), and Arts and Humanities Research Council (AHRC). The authors would like to thank Mr. Ken Munro of the University of St Andrews for programming the mood measures and Sarah Forster, Alan Godfrey, and Simon Halsey for their contributions to the broader NANA project.

References

Allan, C.E., Valkanova, V., Ebmeier, K.P., 2014. Depression in older people is underdiagnosed. The Practitioner. 258(1771), pp.19–22, 2–3.

American Psychiatric Association, 2013. Diagnostic and Statistical Manual of Mental Disorders (5th Ed.). American Psychiatric Publishing, Arlington.

Astell, A.J., Hwang, F., Brown, L.J.E., Timon, C., MacLean, L.M., Smith, T., Adlam, T., Khadra, H., Williams, E.A., , 2014. Validation of the NANA (Novel Assessment of Nutrition and Ageing)

touch screen system for use at home by older adults. Exp. Gerontol. 60, pp.100–107.

Breiman L., Friedman J.H., Olshen R.A., Stone C.J., 1984. Classification and regression trees. Wadsworth & Brooks, Monterey, CA.

Brown, L.J.E., Adlam, T., Hwang, F., Khadra, H., Maclean, L.M., Rudd, B., Smith, T., Timon, C., Williams, E.A., Astell, A.J., 2016. Computer-based tools for assessing micro-longitudinal patterns of cognitive function in older adults. *Age*. DOI 10.1007/s11357-016-9934-x

Brown, L.J.E., Adlam, T., Hwang, F., Khadra, H., MacLean, L.M., Rudd, B., Smith, T., Timon, C., Williams, E.A., Astell, A.J. 2016. Computerized Self-Administered Measures of Mood and Appetite for Older Adults: The Novel Assessment of Nutrition and Ageing Toolkit. *J. Appl. Gerontoll*. DOI: 10.1177/0733464816630636

Brown, L.J.E. & Astell, A.J., 2012. Assessing mood in older adults: a conceptual review of methods and approaches. *International psychogeriatrics / IPA*, 24(8), pp.1197–206.

Cain, A.E., Depp, C.A.,Jeste, D. V., 2009. Ecological momentary assessment in aging research: A critical review. *Journal of Psychiatric Research*, 43(11), pp.987–996.

Hastie T., Tibshirani R., Friedman J., 2009. The elements of statistical learning. Springer, New York.

Kaye, J.A., Maxwell, S.A., Mattek, N., Hayes, T.L., Dodge, H., Pavel, M., Jimison, H.B., Wild, K., Boise, L., Zitzelberger, T.A. 2011. Intelligent Systems For Assessing Aging Changes: home-based, unobtrusive, and continuous assessment of aging. J. Gerontol. B Psychol. Sci. Soc.Sci, 66 Suppl 1, pp.i180-90.

Krstajic, D., Buturovic, L. J., Leahy, D. E., Thomas, S. 2014. Cross-validation pitfalls when selecting and assessing regression and classification models. J. Cheminf., 6(1), 10. http://doi.org/10.1186/1758-2946-6-10;

Préville, M. Mechakra Tahiri, S.D., Vasiliadis, H., Quesnel, L., Gontijo-Guerra, S., Lamoureux-Lamarche, C., Berbiche, D., 2015. Association between perceived social stigma against mental

disorders and use of health services for psychological distress symptoms in the older adult population: validity of the STIG scale. Aging & mental health, 19(5), pp.464–74.

Rodriguez, J.D., Perez, A., Lozano, J.A., 2010. Sensitivity Analysis of k-Fold Cross Validation in Prediction Error Estimation. IEEE Transactions on Pattern Analysis and Machine Intelligence, 32(3), pp.569–575.

Tibshirani, R., 1996. Regression Shrinkage and Selection via the Lasso. J Royal Stat Society. Series B (Methodol.), 58(1), pp.267–288.

WHO, 2010. mhGAP intervention guide for mental, neurological and substance use disorders in non-specialized health settings: Mental Health Gap Action Programme (mhGAP)., Available at: http://www.who.int/mental_health/publications/mhGAP_intervention_guide/en/. [Accessed July 15, 2016].

WHO, 2016. Depression Fact Sheet. World Health Organisation Media Centre. Available at: http://www.who.int/mediacentre/factsheets/fs369/en/ [Accessed July 15, 2016].

Yesavage, J.A., Brink, T.L., Rose, T.L., Lum, O., Huang, V., Adey, M. Leirer, V.O., 1982. Development and validation of a geriatric depression screening scale: A preliminary report. J. Psychiatr. Res., 17(1), pp.37–49.

Yesavage, J.A. & Sheikh, J.I., 1986. Geriatric Depression Scale (GDS). Clin. Gerontol., 5(1–2), pp.165–173.

Figure legends

Fig. 1. Graph showing deviance for multiple values of lambda in the LASSO logistic regression procedure. The green dotted line indicates the value for lambda with lowest deviance, while the blue dotted line indicates the value of lambda with lowest deviance plus one standard error.

Fig. 2. Graph showing coefficient model produced by the LASSO logistic regression procedure. The green dotted line indicates the value for lambda with lowest deviance, while the blue dotted line indicates the value of lambda with lowest deviance plus one standard error. Solid coloured lines represent variable coefficients changing as a function of lambda.

Fig. 3. Receiver operating characteristic curve for model produced with LASSO with logistic regression. The area under the curve (AUC) is 0.88, indicating good predictive ability for this model.