Transparent Modelling of Risk Assessment factors for Assisted Living

P. Vance[1], S.A. Coleman[1], D. Kerr[1], L. Lindsay[2], B.J. Taylor[2], E.P. Kerr[1], B. Gardiner[1], T.M. McGinnity[3], C. Wu[4].

[1]School of Computing and Intelligent Systems, Ulster University, N. Ireland, UK.
[2]Institute for Research in Social Sciences, Ulster University, N. Ireland, UK.
[3]School of Science and Technology, Nottingham Trent University, UK.
[4]College of Information Science and Engineering, Northeastern University, Shenyang, China.

email: {p.vance; sa.coleman; d.kerr; lindsay-l; bj.taylor; ep.kerr; b.gardiner}@ulster.ac.uk, martin.mcginnity@ntu.ac.uk, wuchengdong@ise.neu.edu.cn.

***Abstract*—Advances in healthcare and improvements in lifestyle have contributed to a rising population of an aging society. Within the social care profession, this causes concern as resources are continually spread thin resulting in increased stress for individual workers and increased financial implications for established healthcare providers. One possible solution to alleviate stress and free up resources is to employ the use of computational intelligence within a home environment to determine important risk factors which should allow health- and social-care professionals to put preventative measures in place to protect elderly people from harm thereby reducing the financial implications of hospital care. A major limitation of computational risk models can that they can be quite complex and cumbersome due to the richness of data used to derive the model, not all of which is particularly useful for determining associated risk. In this paper, a transparent risk modelling method is presented which as well as computing an overall risk level, also reduces model complexity by determining which data are relevant. The transparent nature of the model allows users to understand the model structure used to compute the risk level which is important if health care professional are expected to use such algorithms in the future.**

Keywords—risk-analysis, computational modelling.

I. INTRODUCTION

Potential risks related to the ageing population are major concerns within the health and social care profession. Risks can be viewed as either positive or negative. For example, families of adults with dementia are familiar with the results of negative risks like falling [1], burns [2], driving, wandering and mismanagement of medication. Positive risks, on the other hand, can be viewed as opportunities to improve; this may include, for example, people who suffer from isolation risk stepping outside their comfort zone by venturing out to meet new people. In general, four classes of risks to elderly people have been identified (See Figure 1 for examples): behavioural (such as isolation, self-esteem, paranoia, etc); personal safety (such as road safety, wandering, personal injury etc.); health style (including, for example, nutrition, heating and hygiene) and financial (including managing bill, money management and financial exploitation).

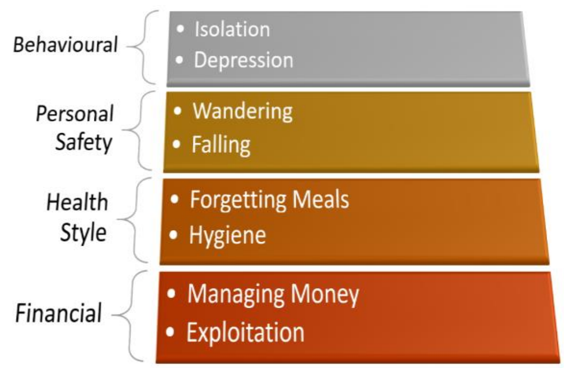

Figure 1: Examples of types of risks linked with an ageing population

Estimating the likelihood of possible harm and working with the individual and family members to manage these risks is an everyday task facing health and social care professionals, particularly social workers. One way of alleviating this burden, especially in times where a growing ageing society is becoming a reality [3], is to utilise computational intelligence to model risk level and thus enable health care professionals to prevent, rather than react to, harmful events. Computational intelligence is a key aspect of computer science, and has received much interest in the recent manipulation of Big Data, which has found application in many research areas, significantly in healthcare. In one case, computational intelligence was used in the diagnosis of breast cancer [4] where a support vector machine (SVM) has been implemented to separate tumours and accurately classify tumour types. This approach has been shown to reduce the diagnosis time without loss of accuracy.

Applying computational intelligence to risk assessment for ageing-in-place would reduce the time and resources required by individual care workers thereby allowing more personal

XXX-X-XXXX-XXXX-X/XX/$XX.00 ©20XX IEEE

attention to be focused where required. To enable this, a process known as activity monitoring [5] can be employed which utilises many discreet sensors within the home such as PIR sensors, pressure mat sensors for under seats, mattress sensors and sensors for environmental aspects such as light and humidity. With these types of sensors fitted around the home, data can be collected that may be interpreted as different events such as cooking, eating, sleeping or simply watching TV [6]. As an example, the activity 'sleeping' would involve a pressure sensor beneath the mattress being activated in combination with all PIR sensors outside the room being inactive for an individual dweller.

While events can easily be derived from the activation state of the various sensors, modelling personal safety or behavioural risks can be unnecessarily complex due to the vast amount of data available containing numerous attributes that may or may not be relevant. For this reason, the use of transparent models for this application is vital as it enables visualisation of the algorithmic decision making and therefore illuminates human understanding of the model output. It facilitates the ability to reduce the complexity of the model and thus improves the scalability of the approach by allowing the removal of attributes or variables which are not required to make a decision. Furthermore, in order for professions such as health and social care to accept and understand decisions made by algorithms, such algorithms need to be debuggable and transparent, and therefore offer 'algorithmic accountability'.

Although there are many computational intelligence approaches that could be applied to risk analysis and modelling, this paper will focus on the ability to reduce model complexity and provide a transparent model. In particular the Non-linear AutoRegressive Moving Average with exogenous inputs (NARMAX) approach will be investigated as this enables us to analyse and validate an input-output coupling using sensitivity analysis and potentially determine decision robustness when faced with incomplete knowledge.

The remainder of this paper is organised as follows. Section II describes the dataset considered to demonstrate this approach whilst Section III gives an overview of the NARMAX modelling process. In Section IV, results are presented which demonstrate the benefits of transparent modelling with a concluding summary presented in Section V.

II. Risk Scenario

As a feasibility study, we demonstrate the approach in terms of a determined risk level classification output as a function of inputted attributes. The data used within this paper originates from the financial sector due to its directness in risk-associated decision-making. The dataset connects a decision of credit suitability as part of a loan application process. The reasoning behind choosing this risk related dataset is that is pertains to a process that most people can associate themselves with, thus they should be able to intuitively rationalise which attributes are important and cross-reference this with the output of the model in terms of the retained attributes. Modelling the process in this way will provide evidence that the particular approach would also operate efficiently when considering data variables which are seemingly irrelevant or unintelligible when applied to the healthcare field.

Data within this set were originally collected for 1000 credit applications and organised into a combination of 20 qualitative and numerical attributes [7], each with a number of subcategories, the result being classified into a good (+1) or bad (-1) category. For modelling purposes, the qualitative attributes were converted to numerical representations with subcategories being expanded into individual variables where required. This resulted in 24 variables which are presented in Table 1.

Table 1: Financial Data Risk Attributes

1	Checking Account Balance	9	Property Owned	17	Loan Purpose: Used Car?
2	Loan Duration (Months)	10	Age (Years)	18	Other Debtors/ Guarantors
3	Credit History	11	Other Credit	19	Other Debt
4	Credit Amount	12	Other Credit with this Bank	20	House: Rented/Free
5	Savings Balance	13	No. Dependents	21	House: Owned/Free
6	Time Employed	14	Telephone?	22	Unemployed/ Management
7	Personal Status and Gender	15	Foreign Worker?	23	Unskilled Job/ Management
8	Time in Present Residence	16	Loan Purpose: New Car?	24	Skilled Job/ Management

Each attribute forms part of an overall intuitive consideration carried out by a financial institution in order to categorise applicants as having either good or bad credit risk. Although each of the attributes listed may seem logical and relevant, not all of them may be important from a modelling perspective. These attributes can be considered as variables, or dimensions from a modelling perspective, which can be assessed for importance and where necessary, can be disregarded, reducing the complexity of the risk model, which provides benefits in terms of speed and scalability when considering larger datasets. To do this, transparent modelling in the form of NARMAX can be utilised to reveal the underlying characteristics of the system. For this purpose, the 1000 applications were split 50/50 into Training/Testing datasets.

III. Transparent Modelling

The NARMAX approach is a popular system identification technique with two fundamental aims. The first is to develop a good approximator of input data such that predictions can be made with high accuracy and minimal error. The second key aim, which is important for this approach, and why the NARMAX philosophy was developed [8], is to find the least computationally expensive model and provide insight into the underlying characteristics. This allows one to analyse and understand the rules that represent the underlying system. In doing so, models may be understood by experts and non-experts alike which is of great importance to healthcare professionals when reviewing why an algorithm has made a particular decision (algorithmic accountability).

The NARMAX model, which is a natural extension of the linear ARMAX model [9] can be defined by:

$$y(t) = F[y(t-1), \dots, y(t-n_y), u(t-d), \dots, u(t-n_u), e(t-1), \dots, e(t-ne)] + e(t) \quad (1)$$

which accounts for the combined effects of noise, modelling errors and unmeasured disturbances concerning the inputs and outputs. Here, $u(t)$ is an input vector, $y(t)$ is an output vector, where n_y is the maximum lag on the output vector; $e(t)$ is system noise which is considered bounded and cannot be measured directly and n_u is the max lag on the input vector. $F[.]$, which is an unknown nonlinear function, is typically taken to be a polynomial expansion of the arguments and d is a time delay which is typically set to $d = 1$.

When developing the NARMAX model, the following steps are adhered to [8]: 1) Structure Detection: the determination of the terms within the model; 2) Model Fitting: tune the coefficients; 3) Validation: validating the model with attention to model overfitting and 4) Prediction: output at a future point in time. 5) Analysis: determination of the underlying dynamics of the system.

Determining the structure of the system forms one of the most important parts of this approach. As the structure is typically unknown prior to implementation [10], a number of options exist to approximate the function $F[.]$ which include polynomial, rational and various artificial neural network implementations [9]. In terms of revealing and analysing the underlying properties of the system, the polynomial models offer the most attractive implementation for this work as they provide a compact mathematical model enabling real time transparent decision making.

The method utilised for model reduction is based on the orthogonal least squares approach outlined in [11] which computes the contribution that each potential model term makes to the system output. This is known as the Error Reduction Ratio (ERR) and provides an indication of which terms can be ignored due to their comparatively minor reduction to the mean squared error. Building the system this way, term by term, exposes the significance of each new term added and avoids model overfitting due to an excessive use of time lags or nonlinear function approximations [9] whilst ensuring that the model is as simple as possible and contains good generalisation properties. Model validation then determines if the model is adequate for the task using the approach outlined in [12] which carries out a correlation based model validation.

The NARMAX approach simulates investigative modelling techniques where the important model terms are weighted and then the model is refined by removing less significant terms [11]. The only difference is that in the NARMAX method, the model terms can be identified directly from the data set. The unknown parameters and system noise can then be estimated and accommodated within the model. These procedures are now well established and have been used in many modelling domains [13].

IV. NARMAX Models and Results

Deriving a NARMAX model is an iterative process where the least significant model terms are removed. As described earlier, the inputs to the model consists of 24 variables relating to the status and holdings of individuals applying for credit and the output pertains to the credit risk being classified as either good (1) or bad (-1). The maximum lag considered for n_y and n_u is set at 0 whilst the threshold value for the error reduction ratio (ERR) is set at 0.05. Convergence of the algorithm is detected by monitoring the change within the estimated parameters which is set at $1e^{-5}$ and is typically achieved in ten iterations [11]. The modelling process for this particular dataset took four iterations to complete, removing four model terms within the first iteration and an additional term in the second needing two further iterations to reach convergence. In total, five variables considered unimportant for model accuracy were removed. In this section, results are outlined with particular focus on the derived terms of the model.

A. Linear ARMAX model

An ARMAX model was constructed using the procedure outlined in Section III containing 19 terms, presented in Equation (2):

y(t) =	+0.05016324092698829600...	
	-0.20333075198127554000	* u(n, 1)...
	+0.01278538092886852300	* u(n, 2)...
	-0.07187081143753938600	* u(n, 3)...
	+0.00163405763866556740	* u(n, 4)...
	-0.04484439999448588300	* u(n, 5)...
	-0.05466350645990122900	* u(n, 6)...
	-0.08059488973512765600	* u(n, 7)...
	-0.02687493345156190300	* u(n, 8)...
	+0.05542928071542430100	* u(n, 9)...

-0.05156063888268464700	* u(n, 11)...	
+0.10614968289695714000	* u(n, 12)...	
+0.12143171358846845000	* u(n, 13)...	
-0.17261043514460983000	* u(n, 15)...	
+0.04581383514072093200	* u(n, 16)...	
-0.32227114579413074000	* u(n, 17)...	
+0.45909019241924703000	* u(n, 18)...	
+0.62276116319950758000	* u(n, 19)...	
-0.14077815495543183000	* u(n, 23)...	
-0.09431997066080444200	* u(n, 24)...	(2)

According to Table 1, the terms removed relate to attributes (10, 14, 20 - 22) which define the age of the applicant, whether they have a telephone or not, whether they own or rent their current residence and if they hold a management position within employment. When reviewing each of the terms removed, one can identify the intuitive reasoning for variables which would be disregarded in the determination of a loan decision. The model may also be rewritten in a simpler format to that of Equation (1) as:

$$y(t) = \Theta_1 u_1(t) + \Theta_2 u_2(t) + \Theta_3 u_3(t) + \ldots + \Theta_n u_n(t) + e(t) \qquad (3)$$

where Θ_x describes the estimated parameter, $u_x(t)$ is the input variable and e(t) is the system noise. The additional benefits that this (ARMAX model) provides is an ease of understanding for non-experts such as healthcare professionals.

Computing the accuracy of the model involved using a threshold of zero and classifying the ARMAX model outputs into either 1 or -1 accordingly. This is illustrated in a scatter plot presented in Figure 2. As can be observed from the clustering around the threshold, there is further scope for investigation the effect of a variable threshold which will form part of our future work. In terms of classification accuracy, the model achieves 73.2% for correct decisions which is an adequate approximation to the input data.

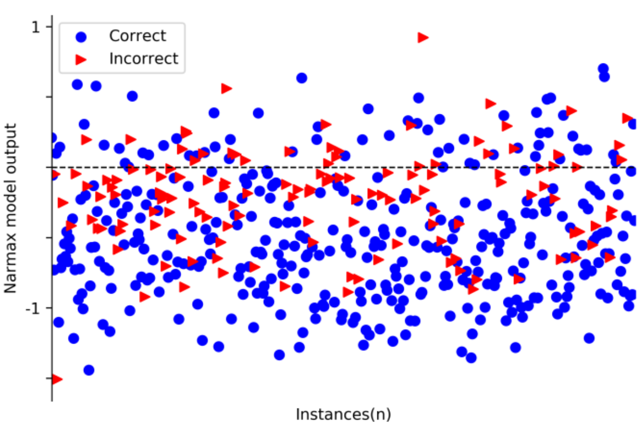

Figure 2: Scatter plot of the linear ARMAX model output.

Removing the indicated variables from the input dataset and repeating the model yielded the same results (73.2% accuracy), confirming the limited effect of the removed variables. The sole difference in model construction was that only two iterations were needed to reach model convergence. Further investigation into the model terms suggested that additional variables could be deleted. Therefore the error reduction ratio (ERR), described in [8] was calculated for each variable to determine which, if any, could be removed to simplify the model further. This ratio provides a simple but effective method of determining a specific terms overall contribution to the models output. Table 2 presents a ranked list of variables according to their relative contribution in reducing the error.

Table 2: Ranked terms according to their ERR values

ERR	Term	
0.000000	+0.0501632409	
11.283843	-0.2033307520	* u1(n)
4.806388	+0.0127853809	* u2(n)
1.112524	+0.6227611632	* u19(n)
0.97182	-0.3222711458	* u17(n)
0.857027	-0.0448444000	* u5(n)
0.683666	+0.0554292807	* u9(n)
0.642018	-0.0805948897	* u7(n)
0.600805	-0.0718708114	* u3(n)
0.532341	+0.0016340576	* u4(n)
0.50663	-0.0546635065	* u6(n)
0.424305	+0.1061496829	* u12(n)
0.331584	+0.4590901924	* u18(n)
0.274131	-0.0515606389	* u11(n)
0.217886	+0.0458138351	* u16(n)
0.135846	-0.0943199707	* u24(n)
0.118993	-0.1726104351	* u15(n)
0.073583	-0.1407781550	* u23(n)
0.061533	+0.1214317136	* u13(n)
0.057918	-0.0268749335	* u8(n)

From the table, three variables with the least contribution relate to an applicant's 'time in present residence' (Table 1, 8), 'no. of dependents' (Table 1, 13) and job type (Table 1, 23). Also evident is that high importance is attributed to variables reflecting one's current banking balance, the duration of the loan and whether or not there is other debt to consider, which is intuitively logical. Using this logical approach, further removal of the three least important variables using an ERR threshold of 0.1 results in an increased accuracy to 74.8% for the linear model, meaning that some variables do in fact negatively impact on a model's performance. This demonstrates how the transparent modelling can be utilised in interpreting what is not important and what variables are deemed most important for a particular model to be used in making a decision. To demonstrate the trade-off between model complexity and accuracy, Figure 3 shows the accuracy of the model when plotted against a variable ERR threshold whilst Figure 4 shows the reduction in terms.

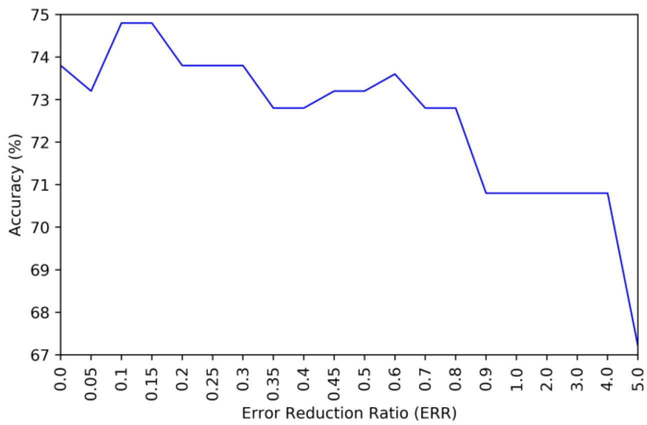

Figure 3: Accuracy vs. variable ERR value

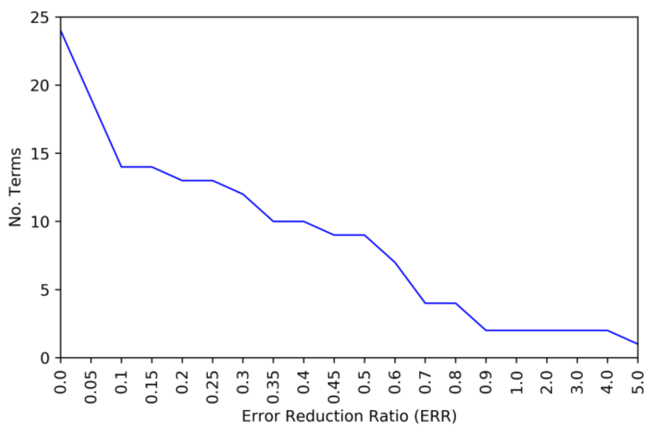

Figure 4: No. Terms vs. variable ERR value

It is also important to note that whilst the ERR values and parameter values do change slightly when computing models with fewer terms, the order of importance does not, therefore Table 2 may be referred to ascertain which terms still remain within the model. When reviewing each of these figures it is evident that using only the current banking balance (Table 1, 1), results in an ERR threshold value of 5.0, achieving a classification accuracy of 67.2% which confirms the importance of this attribute as it is ranked top when referring to its ERR value in Table 2. This is also intuitively logical. Another important observation is that an ERR threshold value of 0.1 yields the best accuracy of the model at 74.8% which results from the contribution of 14 model terms. These results demonstrate that one may compromise between accuracy and complexity, choosing a model which best fits their current application. Given the model has now gone through iterations which computes the accuracy from inclusion of all the terms down to only the most significant term, the next step is to increase the model complexity by observing what contributions, if any, that combinations of terms may provide. This is achieved by computing a quadratic NARMAX model.

B. *Quadratic NARMAX model*

Using the same procedures outlined in Section III, a quadratic NARMAX model was constructed consisting initially of 324 terms which was reduced iteratively to 57 model terms over six iterations using an ERR threshold value of 0.05. Rather than displaying all 57 terms, only the combinations of terms are listed, and ranked, according to their computed ERR values in Table 3. The linear terms omitted remain the same as those listed in Table 2.

Table 3: Ranked terms according to their ERR values

ERR	Term		
2.334253	-0.0003625370	* u2(n)	* u4(n)
1.75549	+0.0001874462	* $u4(n)^2$	
1.060208	+0.0065485778	* u1(n)	* u4(n)
0.938766	-0.0356447390	* u2(n)	* u22(n)
0.882754	-0.4185611808	* $u11(n)^2$	
0.862158	-0.1162594939	* u1(n)	* u11(n)
0.684981	+0.1063418423	* $u7(n)^2$	
0.62634	-0.1283770267	* u3(n)	* u12(n)
0.54548	-0.0024397634	* u4(n)	* u5(n)
0.54291	-3.4891588945	* u15(n)	* u22(n)
0.493778	-0.0251093545	* u10(n)	* u19(n)
0.486379	-0.0514096160	* $u8(n)^2$	
0.433883	+0.0002559988	* u2(n)	* u5(n)
0.397097	+1.4766599296	* u6(n)	* u22(n)
0.380552	+0.1627159201	* u1(n)	* u17(n)
0.372915	-0.0342099249	* u3(n)	* u9(n)
0.357986	-0.0064182959	* u4(n)	* u23(n)
0.354087	+0.0268898278	* u1(n)	* u3(n)
0.353537	-0.1995842341	* u21(n)	* u24(n)
0.343571	-0.0119568312	* u1(n)	* u2(n)
0.337942	+0.0479453111	* u1(n)	* u7(n)
0.332902	-0.2030112446	* u11(n)	* u23(n)
0.326559	+0.0089398818	* u2(n)	* u12(n)
0.284386	+0.2126598657	* u6(n)	* u15(n)
0.248369	+0.0007892382	* u2(n)	* u20(n)
0.244027	-0.1020125592	* u3(n)	* u17(n)
0.242506	-0.0471591313	* $u1(n)^2$	
0.217504	-0.0599102326	* u1(n)	* u9(n)
0.184247	-0.3004021199	* u3(n)	* u22(n)
0.183859	+0.0167266137	* $u5(n)^2$	
0.181801	+0.0420011597	* $u6(n)^2$	
0.142544	+0.6103347551	* u7(n)	* u22(n)
0.135557	-0.0092313786	* u1(n)	* u6(n)
0.120759	+0.0337691675	* u3(n)	* u23(n)
0.113957	+0.0096301646	* u1(n)	* u5(n)
0.108544	+1.3455838032	* u13(n)	* u22(n)
0.098411	-0.0498758373	* u6(n)	* u24(n)
0.080129	+0.0140050483	* u2(n)	* u23(n)

What is interesting here is that model terms previously thought to have been of no relevance are included as combinational terms with other attributes. For instance, input u22 which represents whether one is employed in a management capacity is combined with u2 (loan duration) and is ranked (4^{th} in Table 3) above some of the previous linear terms (Table 2) when comparing ERR values. Again, by increasing the ERR threshold, the effects of term reduction can be observed against computational accuracy. This is shown in Figure 5 and Figure 6 respectively.

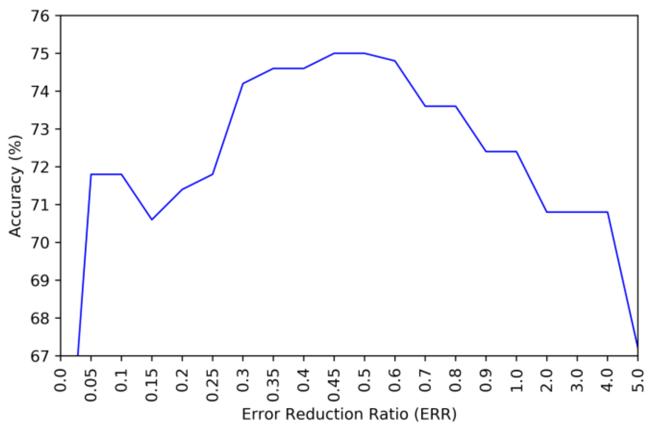

Figure 5: Accuracy vs. variable ERR for the quadratic NARMAX model

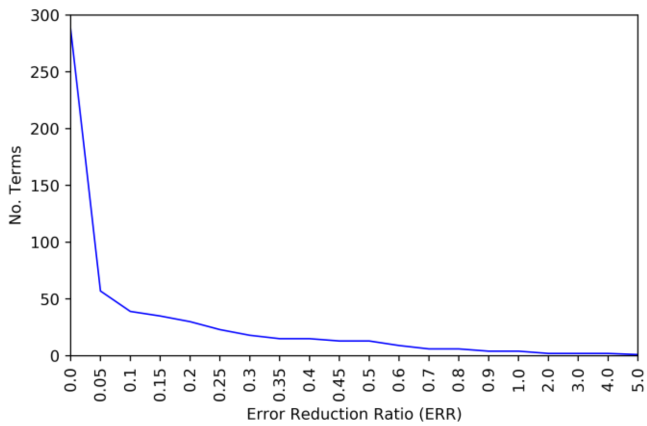

Figure 6: No. Terms vs. variable ERR value for the quadratic NARMAX model

As shown in Figure 5, the accuracy for the quadratic model peaks at 75% which is only 0.2% of an improvement over the linear model. However, the linear model comprised of 14 terms whereas the quadratic model contains one less at 13 due to its ability to combine various attributes. Furthermore, if the accuracy of the quadratic model is matched to that of the linear version to 74.8%, then only 9 terms are required thus reducing the complexity of the model further. On examination of these terms (shown in Table 4), it is evident this is a result of the combination of the loan duration (u2) and credit amount balance (u4) or savings balance (u5) which again is intuitively rational to a non-expert.

Table 4: Ranked terms for 9 term quadratic model

0	0.512251136	
11.28384	-0.1748131891	* u1(n)
4.806388	+0.0413358875	* u2(n)
2.031262	-0.0003439235	* u2(n) * u4(n)
1.712236	+0.0001771754	* $u4(n)^2$
1.202952	-0.2950664922	* u17(n)
0.857027	+0.0392170583	* u5(n)
0.73606	-0.0039976715	* u2(n) * u5(n)
0.642018	-0.6314677578	* u7(n)
0.62426	+0.0917489879	* u9(n)
0.600805	-0.0313954000	* u3(n)
0.58818	+0.1034026864	* $u7(n)^2$
0.532341	-0.0098204761	* u4(n)
0.50663	-0.0583405191	* u6(n)

V. CONCLUSION

Computational modelling results for specific risks regarding elderly people such as falling, hygiene and depression can be difficult for healthcare professionals to accept due to the delicate nature of the problem. In this paper, a method for transparent modelling has been presented that when utilised, provides 'algorithmic accountability' to professionals within the specific area. This would allow healthcare professionals, for example, to understand how a decision has been made by reviewing what data have contributed to the decision, which data variables are most important etc.

The data selected to illustrate the approach originated from the financial sector, chosen for its directness in risk-associated decision-making and enabling of intuitive rationalisation. In selecting this particular dataset, one can see the underlying logic in the risk related decisions and gain an understanding of what is considered most important to the decision-based process. This would be equally transferrable to a healthcare related dataset for those in the healthcare field.

To this end, the work carried out in this paper explores the application and feasibility of modelling specific risks using NARMAX as a method for transparent modelling. We show that by utilising the NARMAX approach, one can intuitively rationalise the terms and discover what the important contributions are in making the decision. In providing this transparency, healthcare professionals can devote more time to eliminating the associated risk factors within one of the associated risk classes which would contribute to prevention of certain risks occurring such as falling, wandering or isolation. If isolation, for example, was deemed an important risk factor that lead to depression, one would concentrate more on communicating with family members or social workers in an effort to eliminate it.

In terms of the computational complexity, we show that the NARMAX model provides a simple approach to determine an efficient model through the ability to prune terms deemed not to contribute effectively. This is achieved primarily through moderation via the ERR value which quantitatively computes the individual contribution that each term affords. Computing the accuracy of such models in this way affords one the option of choosing a model based on the trade-off between complexity and accuracy. This may be particularly beneficial, for instance, for large scale datasets where the deployed model has to operate on limited hardware. Extending this approach to encompass larger datasets within the healthcare field will form the basis of our future work to further improve the efficiency of term selection.

References

[1] L. M. Allan, C. G. Ballard, E. N. Rowan, and R. A. Kenny, "Incidence and prediction of falls in dementia: a prospective study in older people.," *PLoS One*, vol. 4, no. 5, p. e5521, 2009.

[2] J. R. Oyebode, P. Bradley, and J. L. Allen, "Relatives' experiences of frontal-variant frontotemporal dementia.," *Qual Health Res*, vol. 23, no. 2, pp. 156–66, 2013.

[3] P. J. Dobriansky, R. M. Suzman, and R. J. Hodes, "Why population aging matters: A global perspective," *National Institute on Aging, National Institutes of Health, US Department of Health and Human Services, US Department of State*, 2007.

[4] B. Zheng, S. W. Yoon, and S. S. Lam, "Breast cancer diagnosis based on feature extraction using a hybrid of K-means and support vector machine algorithms," *Expert Syst. Appl.*, vol. 41, no. 4, pp. 1476–1482, 2014.

[5] P. Vance, G. Prasad, J. Harkin, and K. Curran, "A wireless approach to device-free localisation (DFL) for indoor environments," in *Assisted Living, IET London*, 2011.

[6] M. Dragone, G. Amato, D. Bacciu, S. Chessa, S. Coleman, M. Di Rocco, C. Gallicchio, C. Gennaro, H. Lozano, L. Maguire, and others, "A cognitive robotic ecology approach to self-configuring and evolving AAL systems," *Engineering Applications of Artificial Intelligence*, vol. 45, pp. 269–280, 2015.

[7] P. Hofmann, "Statlog (German Credit Data) Data Set," *UCI Repository of Machine Learning Databases*, 1994.

[8] S. A. Billings, *Nonlinear system identification: NARMAX methods in the time, frequency, and spatio-temporal domains*. John Wiley \& Sons, 2013.

[9] S. A. Billings and D. Coca, "Identification of NARMAX and related models," *Research Report-University of Sheffield Department of Automatic Control and Systems Engineering*, 2001.

[10] U. Nehmzow, *Robot behaviour: design, description, analysis and modelling*. Springer Science \& Business Media, 2008.

[11] M. Korenberg, S. Billings, Y. Liu, and P. McIlroy, "Orthogonal parameter estimation algorithm for non-linear stochastic systems," *International Journal of Control*, vol. 48, no. 1, pp. 193–210, 1988.

[12] S. Billings and W. Voon, "Correlation based model validity tests for non-linear models," *International journal of Control*, vol. 44, no. 1, pp. 235–244, 1986.

[13] P. J. Vance, G. P. Das, D. Kerr, S. A. Coleman, T. M. McGinnity, T. Gollisch, and J. K. Liu, "Bioinspired Approach to Modelling Retinal Ganglion Cells Using System Identification Techniques," *IEEE Transactions on Neural Networks and Learning Systems*, 2017.