Explainable Artificial Intelligence for Falls Prediction

Anonymous Submission

Abstract. With a rapidly ageing population, it is likely that we will encounter an older adult falling. Falls can cause death, serious injury or harm, loss of confidence and loss of independence. Falling can happen to any of us, however those over 65 years of age can be classified as a group of adults who are more vulnerable and at increased risk of falling. This paper focuses on applying explainable artificial intelligence techniques, in the form of decision trees, to healthcare data in order to predict the risk of falling in older adults. These decision trees could potentially be introduced for health and social care professionals to help aid their judgements when making decisions.

Keywords: Decision Tree, Explainable AI, Classification, Risks, Falls.

1 Introduction

A major public health issue is falls and in particular falls in older adults. A third of the population over 65-years old and half the population over 80-years old are likely to fall at least once per year according to Public Health England [1]. Falling not only affects an older adult physically, it can also lead to an individual having unnecessary stress, loss of confidence and loss of their ability to live independently [2]. If an adult is currently living at home, falls can lead to distress for their families, caregivers and health and social care professionals as they then have to come to a decision if the individual can continue living independently or if other arrangements need to be considered [3]. The future of the individual and their safety and wellbeing is of utmost importance. This means everyone involved needs to communicate effectively with each other about choosing the best outcome for the patient. Health and social care professionals make decisions every day and are focusing more of their attention on risks [4]. In order to help aid their judgements and decisions, decision trees are becoming popular for classifying or calculating risks in healthcare as they can be easily understood and interpreted [5].

The use of Artificial Intelligence (AI) has become popular in industries such as healthcare, education, manufacturing and finance. Explainable AI differs from commonly used opaque AI techniques in that it aims to provide an understanding into how AI decisions are made. Decision tree algorithms provide clarity within machine learning as it is possible to clearly interpret how decisions are reached and the attributes that are deemed important in reaching that decision. Prediction accuracy is used within decision trees to explain how conclusions are made at each stage of the decision-making process, providing a cognitive understanding and ultimately trust from humans.

The remainder of this paper is organized as follows: the motivation to study Explainable AI for Falls Prediction is introduced in Section 2. Section 3 outlines the decision tree results from the study. Finally, Section 4 summarises the work completed in this paper.

2 Explainable AI Methodology for Falls Prediction

Data mining involves extracting useful information from datasets and displaying this in an interpretable way [6]. Decision trees are commonly used for data mining purposes to develop prediction algorithms for a specific target variable. A decision tree can be described as an inverted tree which contains a root node, internal nodes and leaf nodes which are all split into branch-like segments [6]. A decision tree can also be identified as a prediction tree [7]. Decision trees are appealing to use due to their simplicity and their ability to handle mixed data [8]. A decision tree is simply a tree structure that defines a sequence of decisions and their consequences [7]. In this work we use four different types of decision trees to evaluate the effectiveness at measuring the risk of falling: Fast and Frugal Trees, Classification and Regression Trees, Conditional Inference Trees and the J48 decision tree.

Fast and Frugal Trees (FFTs) are a supervised learning algorithm used to create binary classification tasks [9]. This type of decision tree uses sequentially ordered cues, every cue breaks of into two branches, one of these being the exit point. The final cue in the tree will have two exit points for the decision. For the experiments presented here we use the R Studio implementation found under the package FFTrees.

The Classification and Regression Tree (CART) is a form of binary recursive partitioning. Each node in a decision tree can be split into two binary groups. Recursive refers to the binary process being applied over and over again. The partitioning refers to the dataset being split into training and testing sections. An advantage of using the CART decision tree is that it can identify the splitting variables based on searching through all possibilities from the input variables [10]. The building of a tree begins from the root node, this is the beginning of the dataset whereby the variables are split to find the best variable for the root node. The recursive nature of the algorithm ensures that all input variables are checked to find the best variable within the tree. When building the tree, CART recursively splits nodes. As each node is split it is assigned to a predicted class

[10]. Branches are then split below each node in the tree and the decision tree becomes complete when a terminal node is in place as the stopping rule [11]. The CART method used here is implemented in R Studio using the Breiman algorithm [12]. CART uses the 'rpart' method to produce classification decision trees [7]. Rpart follows the simple process of: - 'rpart (formula, data=, method=, control=)' whereby the formula includes: - 'Outcome ~ predictor1 + predictor2' etc. 'Data=' specifies the data frame, 'method=' refers to 'class' if using a classification tree and 'anova' is used for a regression tree. Finally, 'control=' references the optional parameters used for controlling the growth of a tree.

Conditional Inference Tree (CTree) uses two steps to split the tree. CTree determines the variable to be split based on the outcome and the measure of association. After examining all variables, the variable determined to create the best split is then chosen as the root node. Instead of using the Rpart package, CTree uses the Party kit package. By default, the Party kit function uses a quadratic test statistic as it is found to produce more accurate splits [13].

The Waikato Environment for Knowledge Analysis (WEKA) learning environment was used to test the J48 decision tree approach. J48 is an open source Java implementation of the C4.5 algorithm. A J48 decision tree is constructed iteratively, one node at a time. Each lead in the tree represents a classification and the branches that connect the lead to the root node are the conditions that produce the classification. The different transparent decision trees can be compared and evaluated according to their individual predictive accuracies to ensure the model correctly predicts the class of either new or unseen data [14].

3 Experiments and Results

We use the dataset from the Irish Longitudinal Study on Ageing (TILDA) [15]. The dataset comprises information from over eight thousand adults whom are all over the age of 50 and living in the community. The dataset is split into three different waves. Wave 1 incorporates data that were collected between 2009 and 2011. Wave 2 data were collected between 2012 and 2013 and Wave 3 data were collected between 2015 and 2016. TILDA collects data from the community-dwelling participants in waves approximately every two years. In this paper we focus specifically on the use of Wave 1 and Wave 2 data only. The TILDA dataset has been previously used in studies such as predicting the likelihood of recurrent falls in older adults based on previous falls [16].

The four different decision tree methods outlined in Section 2 were used for this study, namely, Fast and Frugal Trees, Classification and Regression Trees, Conditional Inference Trees and a Decision Tree known as J48 in WEKA. The inputs into each of the decision tree algorithms remained the same in all cases. The data were split into a training set and a testing set using a 90:10 split for all four algorithms. The input variables used from the TILDA dataset were the same for all four decision tree algorithms: “Overall Health Description”, “Emotional Mental Health”, “Long-term Health Issues”, “Previous Blackout or Fainting”, “Afraid of Falling” and “Joint Replacements”. The target output was defined as either *fall* or *no falls* using a binary classification represented by 0 (no falls) and 1 (falls).

Figure 1 presents the J48 decision tree produced in WEKA. The tree is significantly deeper than the other algorithms before the terminal nodes are defined. The J48 decision tree produced the best predictive classification accuracy out of all four trees. Each terminal node includes a final outcome and two predictive accuracies. Take for example:

- Long-Term Health Issues – Yes (They do have a long-term health issue)
- Afraid of Falling – Yes (They are afraid of falling)
- Overall Health – Good (Their overall health is good)
- The terminal node concludes with No Falls, the first predictive accuracy is a 0.67 chance of no falls. The second predictive accuracy in the same terminal node is a 0.33 chance of falls.

However, if their long-term health issue is Yes, Afraid of Falling is Yes and their Overall Health is Poor then the terminal node is different from the above. The terminal node in this instance has a final outcome which is fall as there is a higher chance of a fall than no falls in this case. In each terminal node the outcome, either fall or no falls, relates to the first predictive accuracy in the node.

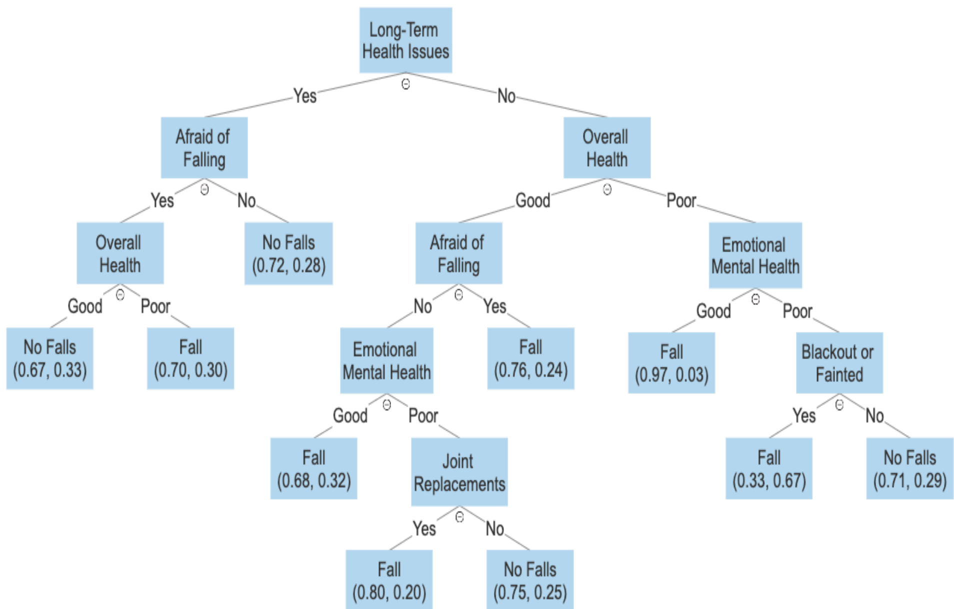

Figure 1 - J48 Decision Tree for Falls and No Falls

Table I explains the J48 rules. The J48 tree is the most complex tree generated in this study. There is a good range of outcomes for falling and not falling. There is a 28% chance of falling if you have a long-term health issue and you are not afraid of falling. If you do have a long-term health issue, you are afraid of falling and have a poor overall health description then you have a higher 70% chance of falling.

TABLE I J48 rules for Falls or No Falls

0.28 Fall when Long-term Health Issues = Yes, Afraid of Falling = No
0.72 No Fall when Long-term Health Issues = Yes, Afraid of Falling = No
0.33 Fall when Long-term Health Issues = Yes, Afraid of Falling = Yes, Overall Health Description = Good
0.67 No Fall when Long-term Health Issues = Yes, Afraid of Falling = Yes, Overall Health Description = Good
0.70 Fall when Long-term Health Issues = Yes, Afraid of Falling = Yes, Overall Health Description = Poor
0.30 No Fall when Long-term Health Issues = Yes, Afraid of Falling = Yes, Overall Health Description = Poor
0.68 Fall when Long-term Health Issues = No, Overall Health Description = Good, Afraid of Falling = No & Emotional Mental Health = Good
0.32 No Fall when Long-term Health Issues = No, Overall Health Description = Good, Afraid of Falling = No & Emotional Mental Health = Good
0.80 Fall when Long-term Health Issues = No, Overall Health Description = Good, Afraid of Falling = No, Emotional Mental Health = Poor & Joint Replacements = Yes
0.20 No Fall when Long-term Health Issues = No, Overall Health Description = Good, Afraid of Falling = No, Emotional Mental Health = Poor & Joint Replacements = Yes
0.25 Fall when Long-term Health Issues = No, Overall Health Description = Good, Afraid of Falling = No, Emotional Mental Health = Poor & Joint Replacements = No
0.75 No Fall when Long-term Health Issues = No, Overall Health Description = Good, Afraid of Falling = No, Emotional Mental Health = Poor & Joint Replacements = No
0.76 Fall when Long-term Health Issues = No, Overall Health Description = Good & Afraid of Falling = Yes
0.24 No Fall when Long-term Health Issues = No, Overall Health Description = Good & Afraid of Falling = Yes
0.97 Fall when Long-term Health Issues = No, Overall Health Description = Poor & Emotional Mental Health = Good

0.03 No Fall when Long-term Health Issues = No, Overall Health Description = Poor & Emotional Mental Health = Good
0.29 Fall when Long-term Health Issues = No, Overall Health Description = Poor, Emotional Mental Health = Poor & Blackout/Fainted = No
0.71 No Fall when Long-term Health Issues = No, Overall Health Description = Poor, Emotional Mental Health = Poor & Blackout/Fainted = No
0.33 Fall when Long-term Health Issues = No, Overall Health Description = Poor, Emotional Mental Health = Poor & Blackout/Fainted = Yes
0.67 No Fall when Long-term Health Issues = No, Overall Health Description = Poor, Emotional Mental Health = Poor & Blackout/Fainted = Yes

In Figure 2 the Fast and Frugal tree is illustrated. In this tree it can be noted that if a subject has no long-term health issues, the tree branches off straight away into a terminal node. The Fast and Frugal Tree predicts that if you have no long-term health issues then you are not likely to have any falls. The J48 decision tree had an accuracy result of 69%, whereas the Fast and Frugal tree correctly classified 67% in the overall predicted accuracy (See Table V). Both of these trees performed the highest out of the four and there is no significant differences between both trees.

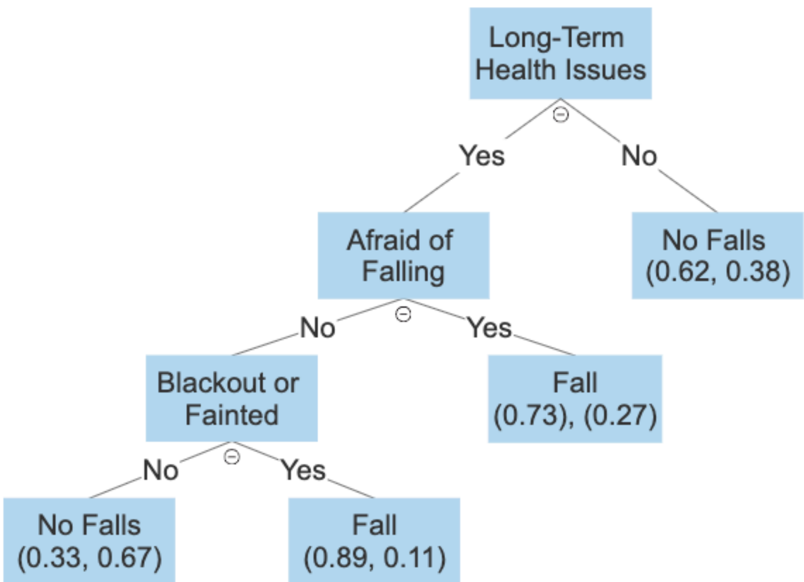

Figure 2 Fast & Frugal Tree of Falls and No Falls

Presented in Table II are the rules that correspond to the Fast and Frugal tree in Figure 2. The probabilities of falls are higher in two out of three of the condition statements. One of the condition statements has a considerably high 89% chance of falling if someone has a long-term health issue and if they are afraid of falling. The probabilities of each terminal node can be found in Table II.

TABLE II: Fast & Frugal Trees (FFTree)

0.33 Fall when Long-term Health Issues = Yes, Afraid of Falling = No & Blackout/Fainted = No
0.67 No Fall when Long-term Health Issues = Yes, Afraid of Falling = No & Blackout/Fainted = No
0.73 Fall when Long-term Health Issues = Yes & Afraid of Falling = Yes
0.27 No Fall when Long-term Health Issues = Yes & Afraid of Falling = Yes
0.89 Fall when Long-term Health Issues = Yes, Afraid of Falling = No & Blackout/Fainted = Yes
0.11 No Fall when Long-term Health Issues = Yes, Afraid of Falling = No & Blackout/Fainted = Yes

Presented in Table III are the rules that were generated by the Conditional Inference Tree. It can be seen from the input variables that "Joint Replacements" were discarded by the algorithm due to having no significance. The highest probability of falls is 59% where someone does not have a long-term health issue, they are not afraid of falling, their overall health is poor and their emotional mental health is poor. The Conditional Inference tree performed poorly in comparison to the Fast and Frugal and J48 decision tree with a classification accuracy of 60%.

TABLE III: Conditional Inference Tree (CTree)

0.59 Fall when Long-term Health Issues = No, Afraid of Falling = No, Overall Health = Poor & Emotional Mental Health = Poor
0.41 No Fall when Long-term Health Issues = No, Afraid of Falling = No, Overall Health = Poor & Emotional Mental Health = Poor
0.48 Fall when Long-term Health Issues = No, Afraid of Falling = No & Overall Health = Good
0.52 No Fall when Long-term Health Issues = No, Afraid of Falling = No & Overall Health = Good
0.46 Fall when Long-term Health Issues = Yes & Afraid of Falling = Yes
0.54 Don't Fall when Long-term Health Issues = Yes & Afraid of Falling = Yes
0.53 Fall when Long-term Health Issues = Yes, Afraid of Falling = No & Blackout/Fainted = Yes

0.47 No Fall when Long-term Health Issues = Yes, Afraid of Falling = No & Black-out/Fainted = Yes

The CART tree algorithm focused only on "Long-term Health issues" and "Afraid of Falling" and its output is presented in Table IV. The probability of falling when someone does have a long-term health issue and when someone is afraid of falling is 54%, this is significantly lower than the Fast and Frugal tree which had 89% for the same circumstances. In previous work, Brighton compared the predictive accuracy of fast and frugal trees with classification and regression trees and found that varying the size of the training sets made a difference to which tree outperformed the other which may explain the difference between the two predictive accuracies [17]. The CART algorithm was the poorest performing tree possibly because it only created a tree using a small number of the risk factors that were inputted.

TABLE IV: Classification & Regression Trees (CART)

0.57 Fall when Long-term Health Issues = Yes
0.43 No Fall when Long-term Health Issues = No
0.54 Fall when Long-term Health Issues = Yes, Afraid of Falling = Yes
0.46 No Fall when Long-term Health Issues = Yes, & Afraid of Falling = No

Using the testing data, the prediction accuracy for each of the decision trees is presented in Table V. The overall differences between each of the algorithms are not significant however the best performing algorithm, Decision Tree (J48) obtained an overall accuracy of 69% correct classifications of fall or no fall.

TABLE V: Results for each Decision Tree

Decision Tree Classifier	Correctly Classified %
Fast & Frugal Trees (FFT)	0.67
Classification & Regression Trees (CART)	0.58
Conditional Inference Trees (CTree)	0.60
Decision Tree (J48)	0.69

Although the results are in favour of the J48 decision tree, Fast and Frugal trees may be preferred by Health and Social Care professionals as a Fast and Frugal tree has two branches at every node and each branch is the opposite of each other. This allows professionals to process the tree much quicker and use the process of elimination while interpreting the tree. Gerd Gigerenzer [18] states that the fast and frugal trees are still being used by physicians as they are easily adapted compared with complex machine learning algorithms [18].

4 Conclusion

This study has explored four decision trees algorithms using data from The Irish Longitudinal Study on Ageing. Each decision tree presents the relationship between each of the inputted different risk factors. The health and social care factors that were explored were "Overall Health Description", "Emotional Mental Health", "Long-term Health Issues", "Blackouts/Fainting", "Afraid of Falling" and "Joint Replacements". For all of the algorithms other than J48 decision trees, Joint Replacements were removed as they are considered to have no significance compared to the input factors towards the risk of falling. Considering the overall accuracies, although each of the trees were between 58% to 69% accurate, these results are based on self-declared qualitative data which would be typical of the accuracies obtained for these type of data. It is apparent from the classification results that the explainable decision trees are easily interpreted. The most important aspect of these models is to ensure health and social care professionals understand and accept the models that may help in their day-to-day work with the ability to help and provide the necessary knowledge that can help guide and support their decisions [3]. Further work will consider a visualization dashboard to compare how risks can be visualized in the real world when health and social care professionals are faced with risks every day in their work.

References

1. Fenton, K. *The human cost of falls - Public health matters* (2014). Retrieved January 29, 2020, from https://publichealthmatters.blog.gov.uk/2014/07/17/the-human-cost-of-falls/
2. National Institute for Health and Care Excellence (2013). *Falls in older people: assessing risk and prevention*. Clinical Guideline Nice 2020. www.nice.org.uk/guidance/cg161.
3. Godolphin, W. (2009). *Shared Decision-Making. Healthcare Quarterly*, *12*(sp), e186–e190. https://doi.org/10.12927/hcq.2009.20947
4. Stevenson, M., McDowell, M. and Taylor, B. (2017) Concepts for communication about risk in dementia care: A review of the literature. Doi: 10.1177/1471301216647542
5. Dillibabu. R & Suresh, K. (2018). *Designing a Machine Learning Based Software Risk Assessment Model Using Naïve Bayes Algorithm*. www.tagajournal.com
6. Song, Y. Y., & Lu, Y. (2015). Decision tree methods: applications for classification and prediction. *Shanghai Archives of Psychiatry*, *27*(2), 130–135. https://doi.org/10.11919/j.issn.1002-0829.215044
7. Sharma, A., & Srivastava, A. (2016). *Understanding Decision Tree Algorithms by using R Programming Language*. 177–182.

8. Su, J., & Zhang, H. (n.d.). *A Fast Decision Tree Learning Algorithm Introduction and Related Work*. Retrieved January 20, 2020, from www.aaai.org
9. Phillips, N. D., Neth, H., Woike, J. K., & Gaissmaier, W. (2017). FFTrees: A toolbox to create, visualize, and evaluate fast-and-frugal decision trees. In *Judgment and Decision Making* (Vol. 12, Issue 4).
10. Lewis, R. J. (n.d.). *An Introduction to Classification and Regression Tree (CART) Analysis*. Retrieved January 9, 2020, from https://www.researchgate.net/publication/240719582
11. Venkatasubramaniam, A., Wolfson, J., Mitchell, N., Barnes, T., Jaka, M., & French, S. (2017). Decision trees in epidemiological research. *Emerg Themes Epidemiol*, *14*, 11. https://doi.org/10.1186/s12982-017-0064-4
12. Speybroeck, N. (2012). Classification and regression trees. *International Journal of Public Health*, *57*(1), 243–246. https://doi.org/10.1007/s00038-011-0315-z
13. Hothorn, T., Hornik, K., Wien, W., & Zeileis, A. (n.d.). *ctree: Conditional Inference Trees*, Vignette R package partykit version 1.1-1, https://CRAN.R-project.org/web/packages/partykit/vignettes/ctree.pdf (2016, accessed 28 January 2020)
14. Zheng, Y., Peng, L., Lei, L., & Junjie, Y. (2005). R-C4.5 decision tree model and its applications to health care dataset. *2005 International Conference on Services Systems and Services Management, Proceedings of ICSSSM'05*, *2*, 1099–1103. https://doi.org/10.1109/ICSSSM.2005.1500165
15. *The Irish Longitudinal Study on Ageing,* Trinity College Dublin, The University of Dublin, from https://tilda.tcd.ie/
16. Lindsay, L., Coleman, S., Taylor, B., Kerr, D., & Moorhead, A. (2019). Using Machine Learning Algorithms to Predict the Likelihood of Recurrent Falls in Older Adults. In *15th International Conference on Machine Learning and Data Mining* (pp. 1-5)
17. Guo, P., Pedryz, W. (2014) Human Centric Decision-Making Models for Social Sciences. In *Springer-Verlag Berlin Heidelberg.* Doi:10.1007/978-3-642-39307-5
18. Wasley, D., Araujo, L. S., Raab, M., & Gigerenzer, G. (2015). *The power of simplicity: a fast-and-frugal heuristics approach to performance science*. https://doi.org/10.3389/fpsyg.2015.01672