Gendered Data in Falls Prediction using Machine Learning

Leeanne Lindsay
Institute for Research in Social Sciences
Ulster University
Northern Ireland, UK
lindsay-l@ulster.ac.uk

Sonya Coleman
Intelligent Systems Research Centre
Ulster University
Northern Ireland, UK
sa.coleman@ulster.ac.uk

Dermot Kerr
Intelligent Systems Research Centre
Ulster University
Northern Ireland, UK
d.kerr@ulster.ac.uk

Brian Taylor
Institute for Research in Social Sciences
Ulster University
Northern Ireland, UK
bj.taylor@ulster.ac.uk

Anne Moorhead
Institute for Nursing and Health Research
Ulster University
Northern Ireland, UK
a.moorhead@ulster.ac.uk

***Abstract*— Adults over the age of 65 years may be considered as a vulnerable population prone to having falls which may have huge consequences. Machine learning is being explored as an approach to understanding better the specific risk factors for falling. However most studies use composite population data rather than including data on male or female gender in the analysis. This study focused on using machine learning models utilizing healthcare data to establish whether gendered data gives a more accurate prediction of falling. Splitting the data into male and female gives slightly higher predictive accuracy, however reducing the size of the dataset is likely to give a lower prediction. Such models could usefully inform health and social care professionals in their daily decision-making with individuals and families about optimal care arrangements.**

Keywords— Machine Learning, Male, Female, Falls, Prediction.

I. INTRODUCTION

Public Health England reported that a third of the population of adults over 65 years old is likely to fall at least once per calendar year [1]. They also reported that half of the population over 80 years old is likely to fall once per year. It is estimated that 255,000 adults over the age of 65 years old who fall are admitted to hospital in England [2]. Similar to Public Health England, the Health Executive in Ireland reported that 30% of adults over 65 years old in Ireland will fall at least once per year [3].

It is quite common to be defined as being in a vulnerable group of adults if you are over 65 years of age as you are unfortunately more prone to falls [4]. There is also an increased prevalence of fear of falling in older adults whether or not they have already fallen [5]. With the increased risk of falling in older adults, there is a tendency to lose confidence when living at home independently. There is also a high possibility of losing mobility due to falls, resulting in further injury. Overall falls affect health and social care services due to the increasing costs that arise [3]. It is estimated that falls cost the NHS in England around £2.3 billion per year [6].

Falls affect not only the individual but may also have implications for the individual's family and decisions by health and social care professionals. After an older person falls, a shared decision between individuals, families and health and social care professionals may be made on whether or not the individual can continue living at home safely [7]. Health and Social Care professionals make these types of decisions every day in order to support the well-being of an older adult. Decisions regarding the elderly are best when they are shared between all parties and ensure an appropriate safe care package is in place [7]. Falls may occur due to factors such as poor eyesight, low blood pressure, medication or cognitive impairment [8]. As humans we all experience positive and negative risks in the world we live in. Generally, older adults encounter more negative risks like burns [9], forgetting to take medication or taking too much medication [10] and falling [11]. Figure 1 illustrates a small number of the categorised risk factors that older adults experience. Figure 1 is based on the health care risk factors that have been based on a review of the literature.

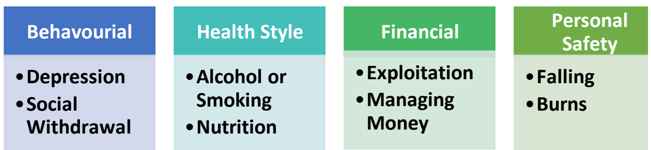

Figure 1 Associated Risks with Older Adults

There is a view that medical research is failing to include gender differences when analyzing data [12]. Therefore, research could potentially be predominantly based on one gender, and generalized over both genders. It is important to include results for both genders in order to implement the best practice for individuals within health and social care research [12]. For example, medical imaging is one field that has implemented gender balance and hence is developing more accurate algorithms in order to help assist medical doctors in diagnosing diseases [13]. Ferrante et al., produced a study which aims to display the importance of gender balance using medical imaging datasets while training artificial intelligence systems for diagnosis [13]. During the study they found that genders that were underrepresented caused a decrease in performance. To overcome bias, training samples should include the details of the sample in terms of male and female so the data can be used for analysis purposes.

In this study we utilize machine learning models to estimate the negative risks of falls in older adults as the study techniques allow us to examine the male and female data individually. We as humans use our attained knowledge to predict different outcomes on a daily basis. Machine learning helps to analyse, design and develop systems [14] with the ability to learn new trends or patterns within the data [15] in a consistent and objective way, but they are subject to underlying bias in the data from which the models are obtained. Machine learning is becoming popular within the health and social care sector to predict adverse outcomes. With the ability to help aid health and social care professionals in decision making, the health sector has adopted machine learning approaches such as diagnosing respiratory conditions from chest x-rays [16], detecting signs of lung cancer at an early stage [17] and identifying patients who need to be moved to ICU [18]. Meaningful information can be extracted with machine learning and used to help assist health and social care professionals in their daily job by finding correlations within data that could ideally be used to help improve the level of care and help reduce costs overall within the health sector [19].

This aim of this study was to use machine learning algorithms to predict the likelihood of falls in older adults by performing experiments and analyses on separate male and female data. The paper is organised as follows: Section 2 outlines machine learning methods and the gendered dataset used; the experiments and results are followed in Section 3; the paper is then brought to a conclusion in Section 4 with future work proposed.

II. Machine Learning & Gendered data

In this study we explore how different machine learning algorithms work in identifying gender differences in the underlying data. The Waikato Environment for Knowledge Analysis (WEKA) [20] is an open source machine learning software programme that has been used for this study. Within WEKA we have used a number of different machine learning approaches as detailed below. Firstly, we use the Naïve Bayes algorithm which implements Bayes Theorem whereby the probability for each class is calculated from the training data supporting both binary and multiclass classification problems. We also use Support Vector Machines with Support Vector Classification (SVC). SVC manages missing data as well as nominal attributes. Multilayer Perceptron is a class of neural networks and contains one or more hidden layers, which we also use within WEKA. WEKA also contains a number of decision tree based approaches. Decision trees are supportive of classification and regression and evaluate data by beginning at the root node of a tree and moving down towards the leaves until a prediction can be made. We use PART which builds a partial C4.5 decision tree in each iteration and the best leaf is then made into a rule. A Random Forest classifier is used which constructs a multitude of trees for classification and regression purposes. Probabilistic approaches such as Bayes Net which is a probabilistic model representing a set of variables and conditional dependencies can also be utilized. Logistic models are also used to predict the probability of a class or event existing by using a logistic function to model binary variables. A Simple Logistic algorithm models the probability of the output in terms of the input. Lastly, Classification via Regression completes classification using regression methods. One regression model is built for each class value. Each of these models will be used in Section III where experiments and results are presented.

In this study we use the Irish Longitudinal Study on Ageing (TILDA) dataset along with the WEKA programme for falls prediction. In particular we are interested in determining if there is any gender bias in these data. The dataset is split into gender based datasets - male and female. We subsequently use the machine learning algorithms to predict the likelihood of falls for males and females independently and thus determine if gender has an effect on machine learning algorithms prediction accuracy and thus if the data have an underlying gender bias. The TILDA dataset is based on adults over the age of 50 years who live in a community dwelling in Ireland [21]. The dataset is split into three different waves each collected in different years. Wave 1 consists of data collected from 2009 to 2011. Wave 2 data were collected during 2012 and 2013 and lastly Wave 3 data were collected in 2015 and 2016. For the purposes of this study, only Wave 1 data have been utilized. TILDA has been used in numerous studies previously [22].

III. Experiments & Results

A number of machine learning algorithms were selected from WEKA's machine learning environment to form predictions alongside the TILDA dataset. These algorithms were used to train models to predict if an individual is likely to fall depending on whether they are male or female. Previous work has identified the risk of an older adult falling using the same dataset [22] but not considering gender differences. This study presents a deeper analysis of the two genders (male and female) to determine any patterns.

There is a number of input factors used in each of the models such as: "Overall health description", "Emotional mental health", "Long-term health issues", "Afraid of falling", "Joint or hip replacements" and any "Previous blackouts or fainting" episodes. A binary classification is the desired target output, corresponding to either fall or no fall. Each input risk factor has been added incrementally to the training inputs of each model to ensure each one is of importance and if not it was subsequently removed. Each dataset was split into a training and testing set and to remain consistent with previous work this split was defined as 90%/10%. To ensure consistency in the reported classifier accuracy, ten-fold cross validation was used throughout all experiments.

Table I displays results based on previous work [22], showing accuracies from 56% - 62%. The lowest accuracy of 56% was obtained using the Multilayer Perceptron machine learning technique and the highest accuracy of 62% was obtained using the Classification via Regression algorithm. The results below were based on using the full dataset *n=3242*.

Table I Machine Learning Algorithm performance using the full dataset (n=3242) [20]

WEKA Classifier	**Correctly Classified %**
Naïve Bayes	61
Support Vector Classification	60
PART	60
Random Forest	57
Decision Tree	59
Bayes Net	61
Logistic	60
Multilayer Perceptron	56
Simple Logistic	60
Classification via Regression	62

In Table II we present results for the same experiment as in Table I using the full dataset except that we have now included gender as another input factor. This enables us to identify if gender has any effect on each of the models and hence determine if there are gender differences.

Table II results demonstrate the significant increase in accuracy when including male and female as a binary input in comparison to Table I. The predictive performance of the algorithms in Table II are higher overall and vary between 57% - 66%. The best performing model is Simple Logistic and Classification via Regression classifying with an accuracy of 66%. The poorest performing algorithm at 57% was again the Multilayer Perceptron. This single model approach has proved to be sufficient in identifying that gender differences exists; the inclusion of gender, enabled all algorithms to classify the data better than when gender was not included (Table I).

Table II Machine Learning Algorithm Performance using the full dataset (n=3242) including male and female

WEKA Classifier	**Correctly Classified %**
Naïve Bayes	64
Support Vector Classification	63
PART	61
Random Forest	58
Decision Tree	63
Bayes Net	64
Logistic	64
Multilayer Perceptron	57
Simple Logistic	66
Classification via Regression	66

As gender differences are evident, we investigated this further by separating the dataset into male and female records, in each dataset *n=1364*. However, this results in a reduced number of records per dataset compared with the results presented in Table I and II. Results using only male data from the dataset are presented in Table III. There are no significant differences in the performance of each of the machine learning algorithms. The highest classification accuracy was consistently achieved by four algorithms; Support Vector Classification, Logistic, Simple Logistic and Classification via Regression which all correctly classify 59% of falls and no falls correctly.

Table III Machine Learning Algorithm Performance using Male data (n=1364)

WEKA Classifier	**Correctly Classified %**
Naïve Bayes	58
Support Vector Classification	59
PART	58
Random Forest	57
Decision Tree	58
Bayes Net	58
Logistic	59
Multilayer Perceptron	57
Simple Logistic	59
Classification via Regression	59

The results for only female data from the dataset are presented in Table IV. The results are again not significantly different in terms of the machine learning algorithms with respect to the best performance. However, we note the overall classification accuracy results are slightly higher than the male only data using the same amount of records as the male dataset *n=1364*. The three algorithms that correctly classified the data with 61% accuracy are Naïve Bayes, Bayes Net and Logistic. The Logistic model remained consistent providing the highest overall accuracy for both the individual male data and individual female data.

Table IV Machine Learning Algorithm Performance using Female data (n=1364)

WEKA Classifier	**Correctly Classified %**
Naïve Bayes	61
Support Vector Classification	60
PART	59
Random Forest	59
Decision Tree	60
Bayes Net	61
Logistic	61
Multilayer Perceptron	59
Simple Logistic	59
Classification via Regression	60

If we compare the results in Table III and Table IV with the results in Table II, we can see that the single model, using the gender data, classifies the output more accurately than two individual gender-based models. However, this could be attributed to the fact that there is more than twice as much data used to generate the results in Table II, compared with Table III and Table IV. Therefore, for fair comparison with balanced datasets, we re-ran the experiment illustrated in Table II, using only 50% of the dataset (*n=1621*) and the results are presented in Table V. This enables direct comparison with the results in Table III and Table IV, and enables a fuller understanding of whether gender is important information and whether inclusion of gender data improves accuracy.

Table V Reduced dataset (n=1621) including male and female using Machine Learning Algorithms

WEKA Classifier	**Correctly Classified %**
Naïve Bayes	60
Support Vector Classification	59
PART	57
Random Forest	56
Decision Tree	60
Bayes Net	60
Logistic	61
Multilayer Perceptron	56
Simple Logistic	61
Classification via Regression	60

A comparison of the results in Table III, Table IV and Table V demonstrates the similarity of results. However the female data in Table IV are slightly better. A comparison of Table V with Table II (the only difference being that Table V uses half the dataset) demonstrates a slight decrease in predictive accuracy due to using fewer records. The use of male and female data as an input variable demonstrates gender differences in the data, and that predictive accuracy can be improved using male and female data as an input variable. The results are not significant enough to justify the use of individual models for gender due to the smaller data set; a single model with gender as an input is sufficient to classify the data. It should be noted that the TILDA dataset is collected through self-declaration and therefore performance accuracy of 50%-70% is as high as one would expect for such a dataset that is not collected in a controlled manner.

IV. Conclusion

This study has explored ten different machine learning algorithms utilizing the data from The Longitudinal Study on Ageing. The risk factors explored were: Overall Health Description, Long-Term Health, Emotional Mental Health, Afraid of Falling, Joint Replacements and Blackouts or Fainting. The dataset was split into two, separating male data from the female data to find correlations or patterns when comparing against the dataset that included both male and female together. This was an attempt to distinguish whether there were any gender differences. A reduction in the size of the dataset lowers predictive accuracy as expected, but splitting the data into male and female gives slightly higher predictive accuracy in both cases with the female data outperforming the male. The slightly higher predictive accuracy of the female compared to the male data suggests that the risk factors used are slightly more relevant for females than males based on this data. For this data it is apparent that separating male and female was beneficial. To be useful in practice and delivery of services, these computer models must be understandable and acceptable to health and social care professionals as potentially they could be of help in their daily when guiding and making decisions by individuals and families. It is important health and social care professionals view males and females differently when looking at risk factors that affect the elderly. Further work proposed is to develop visualization methods to visualize risk to health and social care professionals.

V. References

[1] K. Fenton, "The Human Cost of Falls," 17 July 2014. [Online]. Available: https://publichealthmatters.blog.gov.uk/2014/07/17/the-human-cost-of-falls.

[2] P. H. England, "Public Health Outcomes Framework," 2016. [Online]. Available: www.phoutcomes.info/search/falls.

[3] H. S. Executive, "Falls," NHS Choices, 13 07 2011. [Online]. Available: https://www.hse.ie/eng/health/az/f/falls/. [Accessed 08 07 2020].

[4] B. Johansson, "Balancing integrity vs. risk of falling – nurses' experiences of caring for elderly people with dementia in nursing homes," vol. 1, no. 14, pp. 61-73, 2009.

[5] P. H. England, "Falls and fracture consensus statement: Supporting commissioning for prevention," England, 2017.

[6] N. I. f. H. a. C. Excellence, "Falls in older people: assessing risk and prevention," Clinical Guideline [CG161], England and Wales, 2013.

[7] W. Godolphin, "Shared Decision-Making," *Healthcare Quarterly,* vol. 12, August 2009.

[8] N. Confederation, "Falls Prevention: New approaches to integrated falls prevention services," *Ambulance Service Network Community Health Services Forum,* no. 234, April 2012.

[9] R. J. Oyebode, P. Bradley and L. J. Allen, "Relatives' Experiences of Frontal-Variant Frontotemporal Dementia," vol. 23, no. 2, pp. 156-166, 23 Feb 2013.

[10] C. While, F. Duane, C. Beanland and S. Koch, "Medication management: the perspectives of people with dementia and family carers.," *Dementia (London),* vol. 12, no. 6, pp. 734-750, 2013.

[11] L. Allan, C. Ballard, E. Rowan and R. A. Kenny, "Incidence and Prediction of Falls in Dementia: A Prospective Study in Older People," vol. 4, no. 5, 13 May 2009.

[12] A. Holdcroft, "Gender bias in research:," *Journal of the Royal Society of Medicine,* vol. 100, no. 1, pp. 2-3, January 2007.

[13] A. J. Larrazabala, N. Nieto, V. Peterson, D. H. Milone and E. Ferrante, "Gender imbalance in medical imaging datasets," 2019.

[14] T. Wang, Intelligent Condition Monitoring and Diagnosis System, vol. 93, 2003, p. 132.

[15] P. Parodi, "Computational intelligence techniques for general insurance.," Research dissertation for the Actuarial Profession., 2009.

[16] P. e. a. Rajpurkar, "'CheXNet: Radiologist-level pneumonia detection on chest X-rays with deep learning," 2017. [Online]. Available: https://arxiv.org/abs/1711.05225.

[17] D. Ardila, A. Kiraly and S. e. a. Bharadwaj, "End-to-end lung cancer screening with three-dimensional deep learning on low-dose chest computed tomography," vol. 25, pp. 954-961, 2019.

[18] G. Escobar, B. Turk, A. Ragins, J. Ha, B. Hoberman, S. LeVine, M. Ballesca, V. Liu and P. Kipnis, "Piloting electronic medical record–based early detection of inpatient deterioration in community hospitals," vol. 11, no. 1, 2016.

[19] C. Koh and G. Tan, "Data Mining Applications in Healthcare," vol. 19, no. 2, pp. 64-72, 2005.

[20] E. Frank, M. A. Hall and I. H. Witten, "The WEKA Workbench. Online Appendix for "Data Mining: Practical Machine Learning Tools and Techniques"," 2016. [Online].

[21] T. C. Dublin, "The Irish Longitudinal Study on Ageing," 20 June 2020. [Online]. [Accessed 01 June 2020].

[22] L. Lindsay, S. Coleman, D. Kerr, B. Taylor and A. Moorhead, "Classification of Health Risk Factors to Predict the Risk of Falling," vol. 13, no. 6, 2019.