

Predicting Falls from Operational Data: Insights and Limitations of Using a Non-Specialized Database

Julien Raker¹ ^a, Patrick Elfert¹ ^b, Cletus Brauer² ^c, Marco Eichelberg¹ ^d,
Frerk Muller-von Aschwege¹ ^e and Andreas Hein¹ ^f

¹R&D Division Health, OFFIS - Institute for Information Technology, Oldenburg, Germany

²Johanniter-Unfall-Hilfe e.V., Oldenburg, Germany

fi

Keywords: Fall Prediction, Machine Learning, Health Data, Elderly Care, Predictive Analytics.

Abstract: Falls among the elderly are a significant public health concern. This study investigates the feasibility of predicting falls using an operational dataset from Johanniter-Unfall-Hilfe (JUH) home emergency call system, which was not created under laboratory conditions for scientific purposes. An anonymized dataset containing records from 160,281 participants in Germany was analyzed. Statistical analysis identified 104 out of 400 features significantly associated with falls, though with weak correlations (Cramer's V ranging from 0.006 to 0.071). A one-class Support Vector Machine (SVM) was employed due to the absence of explicit non-fall cases, achieving a true positive rate of 55.10%. The lack of explicit non-fall data prevented evaluation of specificity and overall accuracy. The study demonstrates the potential of using operational datasets for fall prediction but highlights significant limitations due to data quality issues, such as the lack of explicit fall records, absence of non-fall cases, lack of temporal data, and missing values. Recommendations are made to improve data collection practices to enhance the utility of such datasets for predictive modeling.

1 INTRODUCTION

The global demographic shift towards an aging population presents significant challenges for healthcare systems (Nicholas and Smith, 2006). With advances in medicine and public health, people are living longer, but this increased longevity often comes with a higher prevalence of chronic conditions and age-related impairments. One of the most significant risks associated with aging is the increased likelihood of falls (Comans et al., 2013). Falls remain one of the leading cause of injury and morbidity in individuals aged 65 and older, often resulting in severe consequences such as fractures, hospitalization, and loss of independence (Comans et al., 2013; Pfortmueller et al., 2014).


Preventing falls is thus a critical component in improving the quality of life for older adults and


reducing the economic burden on healthcare systems (Becker and Rapp, 2011; Walther et al., 2008). Traditional preventive strategies include physical therapy paired with home modifications, Checking for incompatibilities or unwanted interactions with the current medication or managing fear of falling (Zeeh et al., 2017; Becker and Rapp, 2011; Jansen et al., 2021).


In recent years, machine learning (ML) including artificial intelligence (AI) have emerged as promising tools for enhancing preventive healthcare efforts (Vuppapapati et al., 2019). Through predictive analytics, AI models can analyze large datasets to identify risk factors and early warning signs of health events such as falls, potentially enabling timely interventions (Chattu, 2021).


Numerous studies have explored the use of ML for fall prediction, utilizing various datasets and methodologies. Most of these studies have relied on data collected from sensors, which can be broadly categorized into wearable and non-wearable devices. Wearable devices often include inertial measurement units (IMUs) that track movement and detect anomalies indicative of a fall (El-Bendary et al., 2013; Usmani et al., 2021). Approximately two-


^a  <https://orcid.org/0009-0005-5153-906X>

^b  <https://orcid.org/0000-0002-9834-0702>

^c  <https://orcid.org/0000-0002-2924-7934>

^d  <https://orcid.org/0000-0002-8590-3318>

^e  <https://orcid.org/0009-0001-4960-4097>

^f  <https://orcid.org/0000-0001-8846-2282>

thirds of these studies initially collected data under controlled laboratory conditions to ensure data quality and reliability (Usmani et al., 2021). A notable study using non-sensor data was conducted by Marschollek et al., who aimed to predict whether a patient would fall during an inpatient stay in a geriatric center based on medical records (Marschollek et al., 2012). They analyzed the medical records of over 5,000 patients over a period of 1.5 years. Using the C4.5 decision tree algorithm, trained on the binary attribute of fall or no fall, the model achieved an accuracy of 66%, with a sensitivity of 55.4% and a specificity of 67.1%. While their dataset included results from various geriatric assessments, which significantly influenced the outcomes, such assessments are cost-intensive and not routinely performed, limiting the generalizability of the model. In contrast, general patient-related data such as age, gender, pre-existing conditions (such as chronic diseases or walking impairments), and medication are more commonly available and have also been shown to influence the occurrence of falls (Jansenberger, 2011). However, there is a lack of studies that utilize such general data for fall prediction without relying on specialized assessments or sensor data. This gap highlights the need to explore whether more readily available patient data can be leveraged to predict fall risk, potentially making predictive models more accessible and scalable in real-world settings.

In this study, we aimed to address this gap by exploring the potential of using an existing, non-research-specific database from the Johanniter-Unfall-Hilfe (JUH) home emergency call system to predict falls. The JUH system provides emergency support to elderly or disabled individuals, allowing them to call for help in case of incidents, many of which involve falls (Jörg Lüssem, Thomas Mähner, Hubertus v. Puttkamer, 2021). The database was originally designed for operational purpose, not for research, and thus lacks explicit fall records and comprehensive temporal data. The data includes personal information such as age, medication and previous illnesses, which were recorded once when the home emergency call system was activated (more details in section 2.1). Fall-related incidents were inferred from emergency call records, where falls were often cited as the reason for the call, introducing several limitations in the dataset's applicability for predictive modeling.

We aim to assess whether a ML model, specifically a one-class Support Vector Machine (SVM), can be used to predict falls based on the available data. The one-class SVM was selected due to the dataset's imbalanced nature, containing only fall-related cases and no explicit records of

individuals who have not fallen. This exploratory study seeks to evaluate the feasibility of leveraging a general-purpose dataset for predictive healthcare and to identify key limitations, particularly those inherent to this type of dataset, that future studies must address.

2 METHODS

2.1 Data Source

The data used in this study originates from the JUH home emergency call system, a service designed to assist elderly or disabled individuals by providing immediate emergency support (Jörg Lüssem, Thomas Mähner, Hubertus v. Puttkamer, 2021). The dataset spans from January 1, 2012, to October 16, 2020, and includes anonymized records of participants located in Lower Saxony and Bremen, Germany.

2.1.1 Data Characteristics

The initial dataset comprised records from 160,281 participants, with an average age of 82.73 years (± 9 years) and a predominance of female participants (71.84%). The dataset included:

- **Demographic Information:** Age, gender, and living conditions.
- **Medical History:** Chronic conditions, prior illnesses, and medications.
- **Emergency Call Records:** Details of emergency calls, including timestamps and free-text comments describing the nature of the incident.
- **Care Level:** Classified according to the German care insurance system, indicating the degree of care required by the individual (Nadash et al., 2018).

This is static data that was recorded once when the patient joined the home emergency call system, with the exception of the degree of care, which was updated depending on the status.

2.1.2 Data Limitations

Fall incidents were not systematically documented but had to be inferred from free-text comments in the emergency call logs, which were recorded by the operator during or after the call. Consequently, explicit records were also unavailable for individuals who did not experience falls, as the information on falls was derived solely from emergency calls. This introduces a challenge, as people may fall

without contacting emergency services, complicating the creation of a reliable control group.

2.2 Data Preprocessing

The following preprocessing steps were undertaken to prepare the data for analysis:

1. Extraction of Fall Incidents:

- A keyword search on the free-text comments was applied in the emergency call records to identify potential fall incidents. Keywords included different terms of the word „fall“ in German (e.g. „gefallen“, „gestürzt“, „hingefallen“).
- To increase accuracy, the text has been converted to lower case and umlauts have been removed.

2. Handling Missing Values:

- Due to the significant proportion of missing data, with only 33.57% of participants having complete data available, individuals with incomplete datasets were excluded from the analysis.

3. Feature Selection:

- The Pearson chi-square test of independence was used to identify features significantly associated with fall incidents (see section 2.3).

4. Encoding Categorical Variables:

- Categorical variables were converted into numerical format using one-hot encoding to make them suitable for ML algorithms.

5. Feature Scaling:

- Numerical features were standardized to ensure that all features contribute equally to the model training process.

6. Data Splitting:

- The dataset was split into training and testing sets using a 70/30 ratio.

2.3 Statistical Analysis

To identify features significantly associated with fall incidents, a Pearson chi-square test was applied to the dataset. This test evaluates the relationship between categorical variables (e.g., age group, chronic conditions) and the likelihood of a fall by comparing observed frequencies of falls within one year with expected frequencies under the assumption of independence. The strength of these associations was quantified using Cramer's V, to provide insights

into which feature were most relevant to fall prediction.

2.4 Predictive Modeling

Since the dataset contains only one class, the selection of suitable models is limited. We used algorithms from the field of outlier or anomaly detection for this purpose. Examples of this are one-class SVMs, Isolation Forest or Local Outlier Factor.

Given that one-class SVMs are well-suited for problems where the goal is to identify patterns in datasets containing only positive instances (An et al., 2015), they were chosen for this task. The participants were divided into two groups. Those who made one or more calls due to falls within one year, and those who made no fall-related call within that year. To ensure all features have values in the same range, the data was scaled using the StandardScaler by scikit-learn. Furthermore, the dataset was split into training and test sets, with the training exclusively conducted on the class of participants who experienced falls, and the testing carried out on all data.

2.4.1 Model Evaluation

Due to the absence of non-fall records, model evaluation was conducted by focusing on the models ability to correctly identify actual fall instances within the test set. The following metrics were used to evaluate model performance:

- **True Positive Rate (Sensitivity):** Proportion of actual falls correctly identified by the model.
- **False Negative Rate (Specificity):** Proportion of actual falls incorrectly classified as non-falls.

3 RESULTS

3.1 Descriptive Statistics

For the chi-square test, a total of 146,263 participants were included. The participants had an average age of 83.81 years (± 8.14 years), with ages ranging from 53 to 110. The majority of participants were female (71.75%), while 27.32% were male. 15,262 (10.43%) of those participants are marked as fallen. For 89.57% of participants, no information is available as to whether they have fallen or not.

For the predictive model, only the participants who are marked as fallen and did not have any missing values were included into analysis. The dataset was

thus reduced to a total of 6038 participants, with an average age of 84.65 (± 7.39 years), ranged from 53 to 105. Here too, women were more frequently represented (63.23%) than men (36.40%).

3.2 Statistical Associations

The Pearson chi-square test was applied to identify features significantly associated with fall incidents. Among the 400 features analyzed, 104 demonstrated a statistically significant association with falls ($p < 0.05$).

The strength of these associations was quantified using Cramer's V, with values ranging from 0.006 to 0.071, indicating a very weak correlation. The 20 features with the highest correlation can be seen in Table 1.

3.3 Model Performance

A one-class SVM was employed to predict fall incidents within a year based on the significant features identified in the statistical analysis. The OneClassSVM class from scikit-learn was used for this task. A grid search was performed to identify the best parameters. It was found that the default settings with the sigmoid kernel performed best. The following performance metrics were obtained:

- The model achieved a **sensitivity** of 55.10%, meaning it correctly identified 55.10% of individuals who experienced falls.
- The **false negative rate** was 44.90%, indicating the proportion of fall incidents that the model failed to predict.
- The model classified overall 45.39% of the test participants as at risk of falling.

The confusion matrix (shown in Table 2) summarizes the model's performance.

4 DISCUSSION AND CONCLUSION

This study investigated the feasibility of using a non-research-specific, operational dataset from the JUH home emergency call system to predict falls among the elderly using a one-class SVM. The findings highlight both the potential and the limitations of leveraging such datasets for predictive healthcare modeling.

4.1 Interpretation of Results

The one-class SVM achieved a true positive rate (sensitivity) of 55.10%, correctly identifying over half of the fall incidents in the test set. However, a significant proportion of actual falls (44.90%) were not detected by the model. The statistical analysis revealed that 104 out of 400 features were significantly associated with falls ($p < 0.05$), although the strength of these associations was weak (Cramer's V ranging from 0.006 to 0.071). However, it must be considered here that in the Chi-Square test the group of participants who had fallen was compared with the unclear group (not known whether they had fallen or not). The strength of the correlation can therefore only be viewed with caution.

Notably, several of the 20 features with the highest Cramer's V values correspond to known risk factors for falls, such as walking disabilities, obesity and parkinson (see Table 1). Only the back pain factor showed no known connection to falls. Despite the strong limitation of the data and thus the low strength of the correlation, it was possible to identify characteristics that are known to be associated with falls.

4.2 Limitations of Using Operational Data

Several inherent limitations of the JUH dataset impacted the study's outcomes:

- Fall incidents were not explicitly recorded but inferred from free-text comments in the emergency call logs. This method may have led to underreporting, variations in terminology or spelling errors.
- The dataset did not include explicit records of individuals who did not experience falls, preventing the creation of a control group. This limitation restricted the modeling approach to one-class classification and impeded the evaluation of the models specificity and overall accuracy.
- When changes are made to the data, the previous records are simply overwritten, with no trace left of the alterations. This lack of historical data prevents tracking whether and how patient information has been updated over time. Consequently, it becomes challenging to incorporate temporal trends and monitor changes in patients' conditions. Temporal data, however, are essential for identifying patterns that may

Table 1: The 20 characteristics with the highest Cramer's V. For characteristics with a specified source, the correlation has already been established in previous studies. For characteristics without a source, no correlation is known to date.

Feature	Cramers'V	References
Walking disability	0.070856	(Bergland et al., 2003)
Degree of care	0.048084	(Palm, 2024)
Obesity	0.046093	(Fjeldstad et al., 2008)
Parkinson	0.042153	(Kerr et al., 2010)
Living condition	0.041385	(Walther et al., 2008)
Antihypertensive medication	0.040181	(Klein et al., 2013)
Hearing, seeing and walking impairment	0.039203	(Bergland et al., 2003)
Hearing aid	0.036770	(Bergland et al., 2003)
Back pain	0.031729	<i>none</i>
Hearing impairment	0.031387	(Bergland et al., 2003)
Arthrosis	0.030739	(Rodrigues et al., 2014)
Alcohol abuse	0.024808	(Lima et al., 2009)
Diabetes mellitus (insulin-dependent)	0.024274	(Vinik et al., 2017)
Unsteady blood pressure	0.024247	(Klein et al., 2013)
Diabetes mellitus	0.022882	(Vinik et al., 2017)
Gender	0.021822	(Stevens and Sogolow, 2005)
Hearing impaired	0.021373	(Bergland et al., 2003)
Insulin	0.020749	(Vinik et al., 2017)
Circulatory disorders	0.020547	(Jansen et al., 2016)
Apoplexy	0.020485	(Su et al., 2021)

Table 2: Confusion matrix of the one-class classification model.

	Predicted	
	Fallen	Unknown
Actual Fallen	0.5510	0.4490
Actual Unknown	0.4485	0.5514

signal an increased risk of falls and for enhancing the accuracy of predictive models.

- A significant proportion of records had missing values in key features such as medical history and medication. It was unclear whether missing entries indicated the absence of conditions or a lack of data entry. This uncertainty led to the exclusion of most participants from the analysis and potentially introduced bias.

These limitations underscore the challenges of using operational data not originally designed for research purpose. Data structure, completeness, and quality significantly influence the effectiveness of ML models in healthcare applications.

Still, it makes sense to use operational data because it is real data. Recording data under laboratory conditions is time-consuming, costly and does not necessarily reflect the data in reality. Using existing data can therefore be effective.

4.3 Recommendations for Improved Data Collection Practices

To enhance the utility of operational datasets for predictive modeling, the following recommendations are proposed:

- The introduction of a structured field in the emergency call database table in which the operator can choose between the top n call reasons. This change would improve the accuracy and completeness of incident data and facilitate more precise analyses. Additionally, large language models (LLMs) could be employed to classify free-text fields more effectively, providing further insights in categorizing incident reasons.
- Collect information on individuals who have not experienced falls, possibly through periodic questionnaires or assessments. Participants could report falls or near-falls annually, along with updates on health status and medication changes. This data would enable the creation of a control group and allow the use of traditional binary classification models. We are aware that this approach is difficult to implement and entails significant effort, requiring considerable resources and long-term participant commitment.
- Implement mechanisms to retain historical

records with timestamps when updates are made to patients information. Maintaining a temporal record would allow for the analysis of trends and the inclusion of time-dependent features in predictive models, potentially enhancing their accuracy.

- Differentiate between the absence of a condition and missing data by including explicit indicators. For example, use a specific code to denote „no known illness“ versus „data not provided“. This clarification would improve data integrity and allow for more accurate analyses.

Implementing these recommendations might enhance the quality and research utility of operational datasets, facilitating more effective predictive modeling and contributing to improved healthcare outcomes. However, we acknowledge that implementation can be a challenge, as this involves highly sensitive data (e.g. medication, previous illnesses) that requires careful handling of the data. A balanced solution must therefore be found that meets both the need for comprehensive data and the requirement to minimize data in order to protect privacy.

4.4 Contribution to the Field and Future Work

This study contributes to the field of predictive healthcare modeling by illustrating both the potential and challenges of using non-specialized, operational datasets for fall prediction among older adults. The findings highlight the critical importance of data quality, structure, and completeness in developing effective ML models.

Future Work: In the future, collaboration with data providers, such as JUH, to enhance data collection practices can improve the quality of operational datasets. Additionally, exploring advanced modeling techniques, such as semi-supervised learning, may improve the predictive performance. Another promising area is the integration of additional data sources, such as electronic health records or sensor data from wearable devices, to provide a more comprehensive view of patient health and further enhance model accuracy. Finally, conducting longitudinal studies that preserve temporal data will enable a deeper analysis of how changes over time correlate with fall risk, providing valuable insights into long-term patient outcomes.

4.5 Conclusion

In conclusion, while operational datasets like the JUH home emergency call records hold promise for predictive healthcare applications, significant challenges remain due to data limitations. The study demonstrates that, despite certain predictive capability exists, the effectiveness of ML models is heavily dependent on the quality and structure of the underlying data.

Enhancing data collection and management practices is essential to unlock the full potential of such datasets. By implementing the recommended improvements, organizations could transform operational data into more valuable resources for research, ultimately contributing to better healthcare outcomes for the older population.

ACKNOWLEDGMENTS

This work was carried out as part of the project „LivingSmart - Wohnquartiere neu gedacht – Service-gesteuert: lebensnah, integrativ, intelligent, innovativ“ (Elfert et al., 2023) funded by the German Federal Ministry of Education and Research (reference: 02K17A052). We would like to thank the JUH, especially Alexandra Kolozis, for providing the database used in this research.

REFERENCES

- An, W., Liang, M., and Liu, H. (2015). An improved one-class support vector machine classifier for outlier detection. *Proceedings of the institution of mechanical engineers, part c: Journal of mechanical engineering science*, 229(3):580–588.
- Becker, C. and Rapp, K. (2011). Stürze in der geriatric. *Der Internist*, 52(8):939–945.
- Bergland, A., Jarnlo, G.-B., and Laake, K. (2003). Predictors of falls in the elderly by location. *Aging clinical and experimental research*, 15:43–50.
- Chattu, V. K. (2021). A review of artificial intelligence, big data, and blockchain technology applications in medicine and global health. *Big Data and Cognitive Computing*, 5(3):41.
- Comans, T. A., Currin, M. L., Quinn, J., Tippett, V., Rogers, A., and Haines, T. P. (2013). Problems with a great idea: referral by prehospital emergency services to a community-based falls-prevention service. *Injury prevention*, 19(2):134–138.
- El-Bendary, N., Tan, Q., Pivot, F. C., and Lam, A. (2013). Fall detection and prevention for the elderly: A review of trends and challenges. *International Journal on Smart Sensing and Intelligent Systems*, 6(3).

- Elfert, P., Eichelberg, M., and Hein, A. (2023). *LivingSmart: Intelligente Auswerte- und Informationsmodule zur Notfallprädiktion und -prävention für hochaltrige Personen*, pages 21–36. Springer Vieweg, hmd edition.
- Fjeldstad, C., Fjeldstad, A. S., Acree, L. S., Nickel, K. J., and Gardner, A. W. (2008). The influence of obesity on falls and quality of life. *Dynamic Medicine*, 7(1):1–6.
- Jansen, C.-P., Gross, M., Kramer-Gmeiner, F., Blessing, U., Becker, C., and Schwenk, M. (2021). Empfehlungspapier für das körperliche gruppentraining zur sturzprävention bei älteren, zu hause lebenden menschen. *Zeitschrift für Gerontologie und Geriatrie*, pages 1–11.
- Jansen, S., Bhangu, J., de Rooij, S., Daams, J., Kenny, R. A., and van der Velde, N. (2016). The association of cardiovascular disorders and falls: a systematic review. *Journal of the American Medical Directors Association*, 17(3):193–199.
- Jansenberger, H. (2011). *Sturzprävention in Therapie und Training*. Georg Thieme Verlag.
- Jörg Lüssem, Thomas Mähner, Hubertus v. Puttkamer (2021). Aus liebe zum leben: Jahresbericht 2020 der johanniter-unfall-hilfe e.v. https://assets.johanniter.de/JUH/Bundesgeschaeftsstelle/Dokumente/Jahresberichte/2020.Jahresbericht_Johanniter.pdf. (besucht am 26.07.2022).
- Kerr, G. K., Worringham, C. J., Cole, M. H., Lacherez, P. F., Wood, J. M., and Silburn, P. (2010). Predictors of future falls in parkinson disease. *Neurology*, 75(2):116–124.
- Klein, D., Nagel, G., Kleiner, A., Ulmer, H., Rehberger, B., Concin, H., and Rapp, K. (2013). Blood pressure and falls in community-dwelling people aged 60 years and older in the vhm&pp cohort. *BMC geriatrics*, 13:1–7.
- Lima, M. C. P., Simão, M. O., Oliveira, J. B. d., Cavariani, M. B., Tucci, A. M., and Kerr-Correa, F. (2009). Alcohol use and falls among the elderly in metropolitan são paulo, brazil. *Cadernos de saude publica*, 25:2603–2611.
- Marschollek, M., Gövercin, M., Rust, S., Gietzelt, M., Schulze, M., Wolf, K.-H., and Steinhagen-Thiessen, E. (2012). Mining geriatric assessment data for in-patient fall prediction models and high-risk subgroups. *BMC medical informatics and decision making*, 12(1):1–6.
- Nadash, P., Doty, P., and von Schwanenflügel, M. (2018). The german long-term care insurance program: evolution and recent developments. *The Gerontologist*, 58(3):588–597.
- Nicholas, P. K. and Smith, M. F. (2006). Demographic challenges and health in germany. *Population Research and Policy Review*, 25:479–487.
- Palm, R. (2024). Inkontinenz und stürze erschweren die pflege zu hause. *MMW-Fortschritte der Medizin*, 166(3):40–41.
- Pfortmueller, C. A., Kunz, M., Lindner, G., Zisakis, A., Puig, S., and Exadaktylos, A. K. (2014). Fall-related emergency department admission: fall environment and settings and related injury patterns in 6357 patients with special emphasis on the elderly. *The Scientific World Journal*, 2014.
- Rodrigues, I. G., Fraga, G. P., and Barros, M. B. d. A. (2014). Falls among the elderly: risk factors in a population-based study. *Revista Brasileira de Epidemiologia*, 17(3):705–718.
- Stevens, J. A. and Sogolow, E. D. (2005). Gender differences for non-fatal unintentional fall related injuries among older adults. *Injury prevention*, 11(2):115–119.
- Su, Q., Gao, Y., Zhang, J., Tang, J., Song, M., Song, J., Mao, Y., and Pi, H. (2021). Prevalence of fear of falling and its association with physical function and fall history among senior citizens living in rural areas of china. *Frontiers in public health*, 9.
- Usmani, S., Saboor, A., Haris, M., Khan, M. A., and Park, H. (2021). Latest research trends in fall detection and prevention using machine learning: a systematic review. *Sensors*, 21(15):5134.
- Vinik, A. I., Camacho, P., Reddy, S., Valencia, W. M., Trencé, D., Matsumoto, A. M., and Morley, J. E. (2017). Aging, diabetes, and falls. *Endocrine practice*, 23(9):1120–1142.
- Vuppapapati, J. S., Kedari, S., Ilapakurti, A., Vuppapapati, C., Vuppapapati, R., and Kedari, S. (2019). Machine learning infused preventive healthcare for high-risk outpatient elderly. In *Intelligent Systems and Applications: Proceedings of the 2018 Intelligent Systems Conference (IntelliSys) Volume 2*, pages 496–511. Springer.
- Walther, L., Nikolaus, T., Schaaf, H., and Hörmann, K. (2008). Schwindel und stürze im alter. *HNO*, 56(8):833–842.
- Zeeh, J., Reinhardt, Y., and Heppner, H.-J. (2017). Stürze im alter. *MMW-Fortschritte der Medizin*, 159(13):52–58.