Summary

Using a downsampling method with ChatGPT and ML techniques, we obtained a full NEISS dataset across all accidents and age groups from 2013-2022 with six new variables: fall/not fall, prior activity, cause, body position, home location, and facility.

We showed how these new indicators vary by age and sex for the fall-related cases and how they impact the likelihood of falls and post-fall hospitalization of older people. We also revealed seasonal and yearly trends in falls and provided a comparative age-based perspective, underscoring the value of the full ED narratives dataset in studying falls of older adults.

Methodology

First, we preprocessed the narrative field, removing age, sex and most of the diagnosis from the descriptions and expanding the abbreviations.

Second, we labeled the following variables on the samples of narratives using ChatGPT API.

1. **Fall** or not Fall-related accident
2. **Cause** of the accident (*medical such as dizziness, seizures, collapse or not medical)*
3. **Body position** before the accident *(e.g., standing, sitting)*
4. **Activity** before the accident *(e.g., locomotion, manual handling)*
5. **Home location** *(e.g., bathroom, bedroom)* [Only for location = HOME]
6. **Facilities** *(e.g., assisted living, retail spaces)* [Only for location = PUBLIC]

Third, we trained a set of SVM classification models on these samples and predicted coded variables in the full NEISS database (18+ year-olds, 2013-2022). To increase the prediction accuracy, we used ChatGPT to relabel observations with the lowest class probability and retrained the models using the revised sample.

We assessed the pipeline's quality by manually annotating a stratified sample of the predicted categories. The agreement between human coders and ML was high, with accuracies ranging between 80% and 97% by variables. We used imputation methods to predict categories with incomplete text information, reducing bias from inconsistent narrative descriptions.

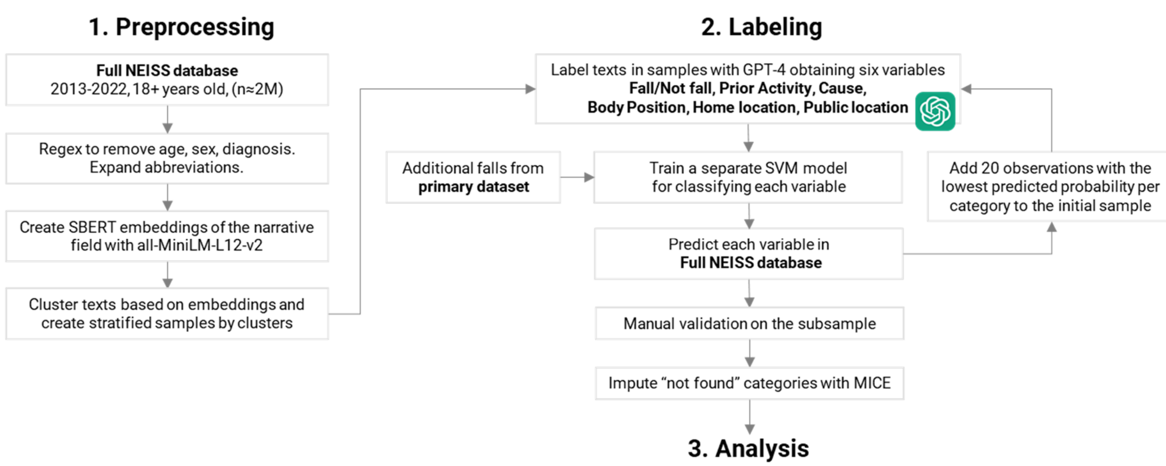

Figure 1. Data Processing Pipeline.

The final database with six predicted variables had 2,046,735 records of all adult accidents during 2013-2022, including 481,268 records of 65+ year-olds with fall-related accidents.

Analysis

Table 1. Descriptive Statistics for the Fall-Related ED Visits of Older Adults, % by column

Variable	Category	Female	Male	65-74	75-84	85+	Total
Cause	Medical	6.6	9.5	8.3	8.2	6.5	7.7
	Not medical	93.4	90.5	91.7	91.8	93.5	92.3
Body position	Standing	75.1	72.7	78.2	75	69.8	74.2
	Sitting	10.5	11.6	8.9	10.5	13	10.9
	Lying	13.8	14.2	11.1	13.6	16.9	13.9
	Other body position	0.7	1.5	1.8	0.9	0.4	1
Activity	Locomotion	51.5	48.5	53.6	51.2	46.7	50.4
	Body positioning	26.8	23.5	21.5	25.1	29.9	25.6
	Stationary	12.5	14.7	11	13.1	15.6	13.3
	Manual handling	4.2	6.2	6	5	4	5
	Sports activity	1.1	2.1	2.4	1.5	0.6	1.5
	Other activity	3.9	5	5.5	4.2	3.2	4.3
Home location	Home in general	18.7	17.2	16.2	19	19.2	18.2
	Bedroom	10.2	9.8	9.1	10.4	10.6	10
	Bathroom	8.7	8.9	8.4	9.1	8.8	8.8
	Stairs	7.2	6.8	9.1	7.5	4.7	7
	Kitchen	3.2	2	2.6	2.9	2.8	2.8
	Living room	2.5	2.6	2.3	2.7	2.6	2.5
	Other home location	12.4	14.3	14.8	13.7	11.1	13.1
Facilities	Assisted living	12	10.1	4.6	8.7	20.1	11.3
	Retail spaces	1.5	0.9	1.5	1.3	1	1.3
	Healthcare facilities	1.2	1.1	1.2	1.2	1.2	1.2
	Food & hospitality	0.6	0.5	0.7	0.6	0.5	0.6
	Entertainment & leisure	0.5	0.6	0.6	0.6	0.4	0.5
	Transportation hubs	0.5	0.5	0.6	0.5	0.3	0.5
	Other type of facility	3.3	3.6	3.6	3.1	3.6	3.4

Table 1 outlines the age and sex distribution for the fall-related ED visits. Over 90% of such visits were due to non-medical reasons. Falls from a standing position were most frequent compared to other positions but decreased with age, while falls from sitting or lying positions rose. Falls were more common during movement, especially in females. While most falls of older adults occurred at home, there were variations in specific locations, with fewer in kitchens and living rooms. Around 11% of falls occurred in assisted living facilities, a number that rose with age.

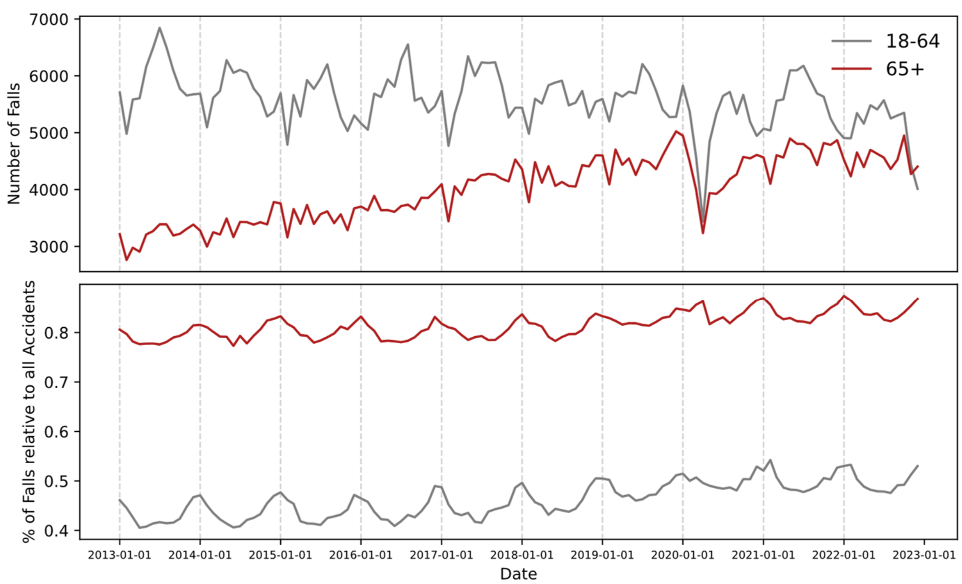

Figure 2. Temporal Trends in Falls by Age Groups.

Figure 2 shows monthly falls for different age groups, suggesting fewer falls in absolute terms for older individuals in winter (upper graph); however, falls relative to other events (bottom) were more common during winter. When adjusting for other incidents, the upward trend in falls in older ages and a COVID-19 drop appeared less pronounced. This may indicate an increased ED use by older individuals due to population ageing rather than a substantial fall-related decline in health.

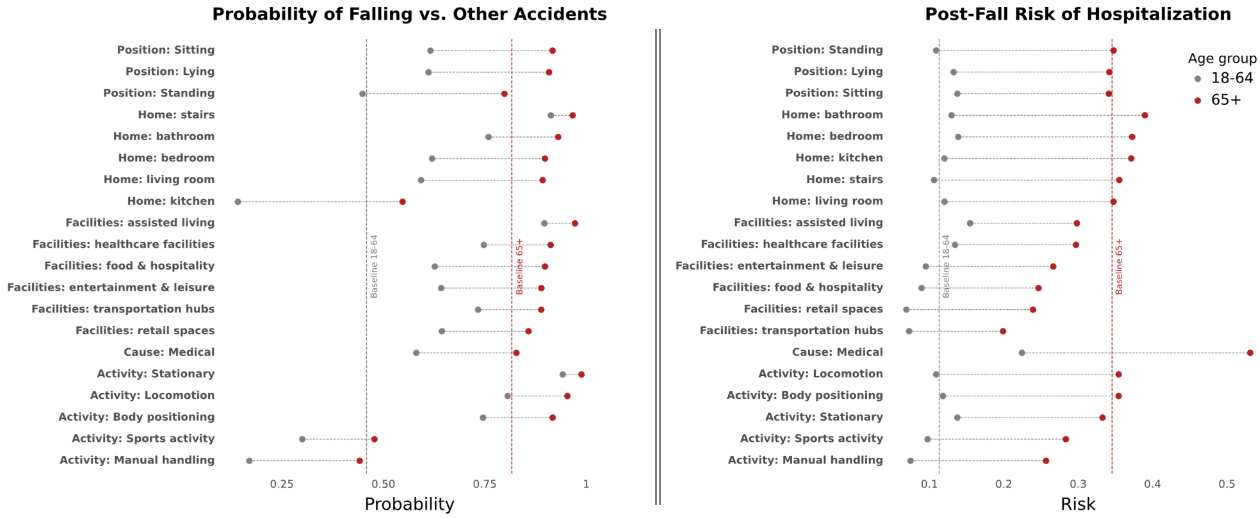

Figure 3. Probability of Falls vs. Other Accidents and Post-Fall Risk of Hospitalization for Older Adults

Figure 3 shows the likelihood of falls vs. other incidents and the risk of post-fall hospitalization. Older individuals were more prone to visit the ED for falls rather than other accidents when sitting or lying, walking on stairs, in assisted living, and stationary. The post-fall hospitalization risks increased due to medical reasons and movement activities. For older people, medical reasons did not lead to more falls compared to non-medical ones. Sports activities and manual tasks reduced fall and post-fall hospitalization chances, implying active people are healthier or more careful. For older individuals, falls in bathrooms or kitchens bring more hospitalizations than in other places. Living rooms and bedrooms posed a higher fall probability for older people than the younger ones; stairs had a similar fall probability for both ages.

Discussion & Conclusion

This study combines the benefits of human-driven categorization with the automation potential of AI. We employed a downsampling strategy, integrating ChatGPT and ML classification methods, to label extensive text samples and extract fall-related variables from the narrative data. The general approach is suitable for diverse research objectives and large datasets, can be applied to other topics (e.g., car accidents or opioid fatalities), and provides broader insights crucial for policymaking.

The limitations of our approach include the need for proprietary ChatGPT API access and manual category creation. Challenges arise from potential ChatGPT inaccuracies, ML model mispredictions, and biases in the narrative data. This study used external validation and missing data imputation to address these biases.

The open-source models like Llama were less effective than GPT-3.5 or GPT-4. GPT-4 showed superior performance over GPT-3.5. Binary variables were more accurately labeled than the multi-categorical ones; Other/Not defined categories were also challenging to classify. Drawing multiple rather than single responses from the GPT API did not enhance the accuracy. Advanced prompt engineering techniques had minimal impact on our results. A simple SVM delivered the best results for the classification phase. Future directions include sensitivity analysis, finetuning the transformer model for classification, and broader testing/validation for other accidents.