Predicting Depression in Old Age: combining Life Course data with Machine Learning

PhD Conference in Social Science

November 24 2022, Luxembourg

Carlotta Montorsi

Alessio Fusco[1] Philippe Van Kerm [1] [2] Stephane Bordas[2]

[1]Luxembourg Institute of Socio-Economic Research (LISER)

[2]University of Luxembourg

Motivation

- **Population ageing** is one of the key challenges of our times. The share of the EU population above the age of 65 is expected to reach almost 25% by 2050 (starting from 19.2% in 2016)

Motivation

- **Population ageing** is one of the key challenges of our times. The share of the EU population above the age of 65 is expected to reach almost 25% by 2050 (starting from 19.2% in 2016)
- Depression in old age is common. In Europe 8.9% of those among 55-64 years old and 8.6% of those 65+ suffer of chronic depression (EUROSTAT, 2019)
- Depression in old age is both under-diagnosed and under-treated in primary care setting
- Depression is an independent predictor of other major diseases: Alzheimer, dementia and diabetes

Objectives

- Prevention strategies and improvements in early identification are essential (WHO, 2016).

Objectives

- Prevention strategies and improvements in early identification are essential (WHO, 2016).
- Predicting depression is a challenge:
    - Complex disease
    - Lack of bio-markers/risk factors
    - Humans subjectivity

Objectives

- Prevention strategies and improvements in early identification are essential (WHO, 2016).
- Predicting depression is a challenge:
    - Complex disease
    - Lack of bio-markers/risk factors
    - Humans subjectivity
- **Supervised Machine Learning Algorithms** may tackle these complexities resulting in high predictive performance

Objectives

- Prevention strategies and improvements in early identification are essential (WHO, 2016).
- Predicting depression is a challenge:
  - Complex disease
  - Lack of bio-markers/risk factors
  - Humans subjectivity
- **Supervised Machine Learning Algorithms** may tackle these complexities resulting in high predictive performance
- Could we preemptively identify clinically depressed individuals from their **past** life histories? Which is the most predictive data configuration?

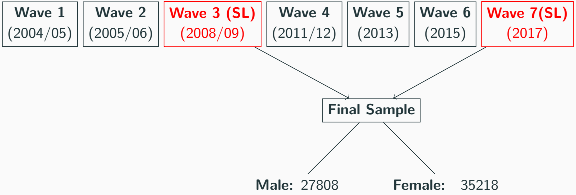
Objectives

- Prevention strategies and improvements in early identification are essential (WHO, 2016).
- Predicting depression is a challenge:
    - Complex disease
    - Lack of bio-markers/risk factors
    - Humans subjectivity
- **Supervised Machine Learning Algorithms** may tackle these complexities resulting in high predictive performance
- Could we preemptively identify clinically depressed individuals from their **past** life histories? Which is the most predictive data configuration?
- Are there differences in life course depressive patterns across genders?

Data

Data Source/Subjects

- The Survey of Health, Ageing and Retirement in Europe (SHARE)
- We drawn **Retrospective information** from SHARELIFE (SL) questionnaire
- Different individuals of wave 3 and wave 7



- We select:
  1. respondents aged < 89 for recall bias
  2. respondents that provide attention during the interview
  3. respondents without missing variables in all depression symptoms across all waves

Measurements framework

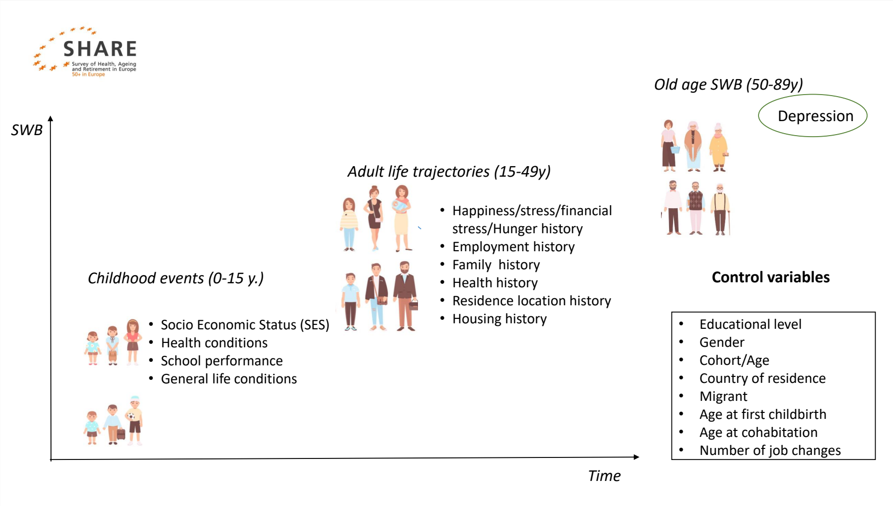


Figure 1: Measurements framework

Depression in SHARE

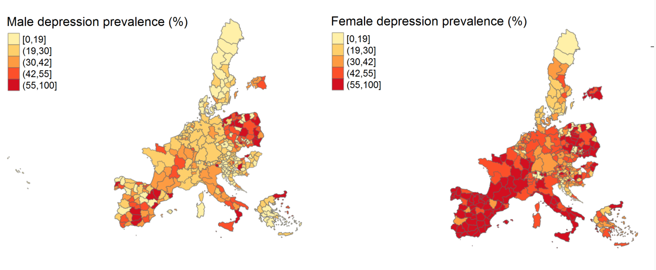


Figure 2: Depression prevalence across genders. Colors represent ventiles of the depression distributions in the pooled sample

- Depression in SHARE is measured by the 12 questions that compose the euro-D instrument: good test-retest reliability and internal consistency (Prince, 1999a)
- Clinical depression threshold: euro-D scale score of higher than 4 is categorized as case of depression (1) and a scale score below four as not depressed (0) (M. Prince et al., 1999b; E. Castro-Costa, M. Dewey, et al.,2008)
- The sample counts 40% individuals with at least one depression measurement in the observation period (46% females, 29% males)

Life Trajectories

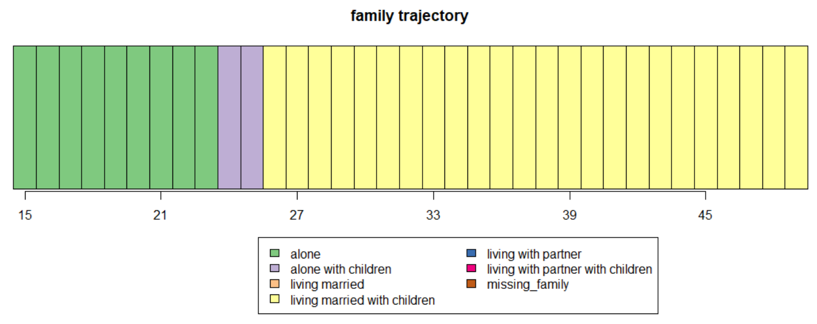
- A life trajectory is defined as the long-term pattern of stability and change, which usually involves multiple transitions. Along this trajectory, each individual may experience many events, either positive or negative

Life Trajectories

- A life trajectory is defined as the long-term pattern of stability and change, which usually involves multiple transitions. Along this trajectory, each individual may experience many events, either positive or negative
- **Sequences** (A. Abbot, 1995):

Life Trajectories

- A life trajectory is defined as the long-term pattern of stability and change, which usually involves multiple transitions. Along this trajectory, each individual may experience many events, either positive or negative
- **Sequences** (A. Abbot, 1995):



Sequences representations

- We construct life trajectories for 6 life dimensions:

  1. Work
  2. Family
  3. Housing arrangement
  4. Location of residence
  5. Health
  6. General life events

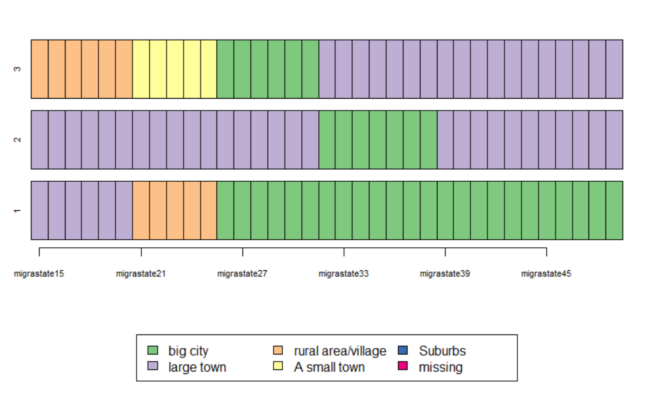
- We operationalize sequence in three different ways:

  1. Clusters or Typologies: distinct groups of individuals' having similar life trajectory ($\sim$ 113 predictors)
  2. Sequences features: timing, duration, sequencing, and entropy ($\sim$ 301 predictors)
  3. Unstructured representation ($\sim$ 306 predictors)

Sequences representations

- We construct life trajectories for 6 life dimensions:
    1. Work
    2. Family
    3. Housing arrangement
    4. Location of residence
    5. Health
    6. General life events
- We operationalize sequence in three different ways:
    1. Clusters or Typologies: distinct groups of individuals' having similar life trajectory (~ 113 predictors) Example cluster
    2. Sequences features: timing, duration, sequencing, and entropy (~ 301 predictors)
    3. Unstructured representation (~ 306 predictors) Example unstructured

Example Features



ID	Duration BC	Duration ST	Duration Rur	LT → BC	LT → Rur → BC	Age(20-25) Rur	Entropy	
1	24	0	5	1	1	1	0.20	...
2	7	1	0	1	0	0	0.12	...
3	6	5	6	0	0	0	0.29	...
...	...	...	...	...	...	...	...	...

Methods

Machine Learning Methods

We explore six different algorithms:

1. Logistic Regression
2. Ridge
3. Lasso
4. Elastic Net
5. Extreme Gradient Boosting
6. Artificial Neural Network

Optimization Routine

Stratified train-test split approach:

1. Training sets: 80% sample; test set: 20% sample
2. Tuning Models: random/grid search + stratified 10-folds cross validation to maximize the Area Under the Precision-Recall Curve (PR-AUC)
3. Compare models' performance on the test set: PR-AUC, AUC, Sensitivity
4. Benchmark model: minimal predictor set with only demographic information, i.e., age, interview year, migrant status, education level, age at first childbirth

Results

Models' PR-AUC Female

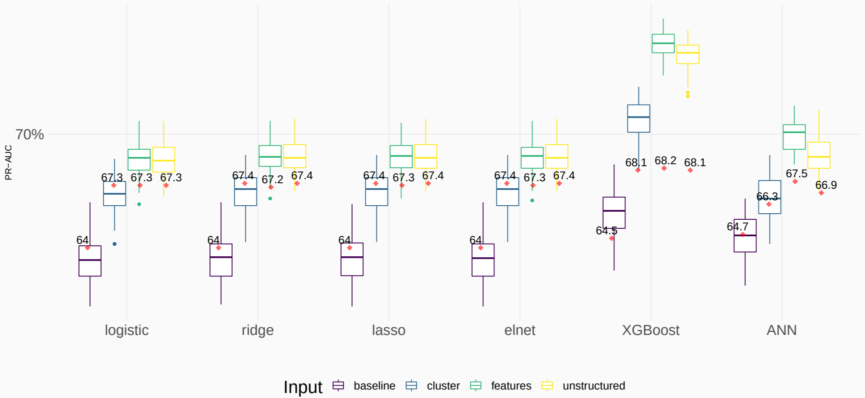


Figure 3: Sensitivity across models and input structures. Performance on the test set. Female sample

- PR-AUC of all algorithms increases along with the increasing dimensionality of the input structure
- We reach a PR-AUC of ~ 68.2% combining the Gradient Boosting with sequential features

Models' PR-AUC Male

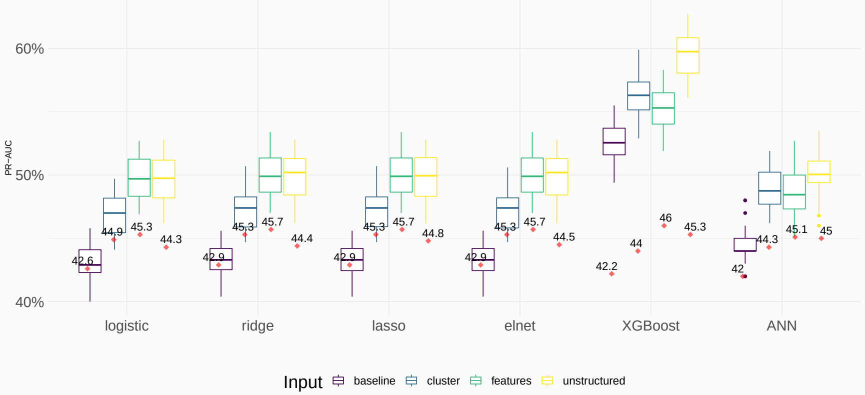


Figure 4: Sensitivity across models and input structures. Performance on the test set. Female sample

- For males, we reach a PR-AUC of ~ 46.2%
- Males' life course trajectories are less informative than females' trajectories.

SHapley Additive ExPlanation (SHAP)

- SHAP values inform on how much each input variable contribute to create the final predicted probability
- Example: we are giving the Gradient Boosting model the life course information of a Slovenian Female of age 59, not depressed

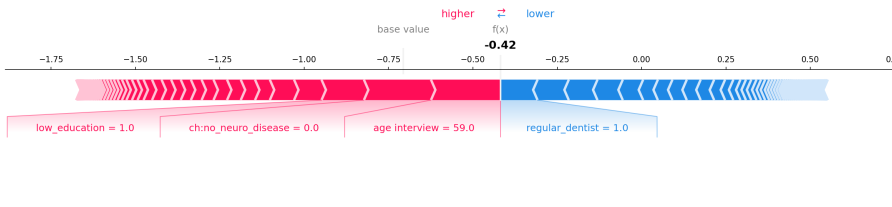


Figure 5: A SHAP force plot of a single individual. In **bold** is the predicted odd ratio, which correspond to 0.39 probability of being depressed. Red represents features that pushed the model probability score higher, blue represents features that pushed the score lower.

Depression Patterns Across Gender

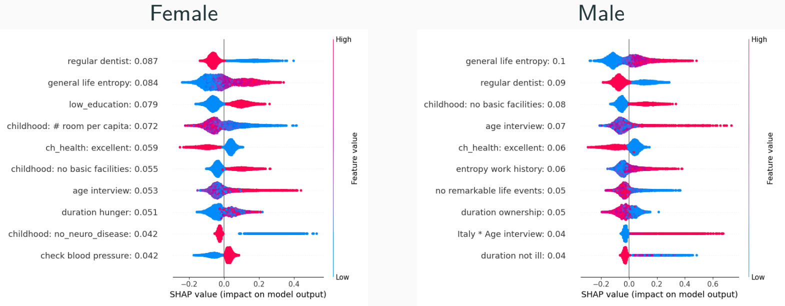


Figure 6: Left is female and right is male. *Note*: Importance of the features in descending order of their importance based on Shapley values. Color for each feature shows the positive or negative correlation with the target outcome.

- For **both genders** and across models: age, fragile health conditions in childhood and adulthood, low education, and low dental care increases depression risk
- Only for **male** and across models: house ownership's duration decreases the risk of depression
- For both genders but only black box methods: higher general life entropy increases the depression risk

Conclusion

- Life histories predict some future clinically depressed individuals but are **not** able to perfectly detect them
- The data required for achieving the highest predictive performance is more complex than what has been traditionally used in well-being studies
- We identify new idiosyncratic and common patterns across genders
- Interpretable machine learning tools may support the hypothesis creation process
- Sub-samples: Performance across thresholds
- Predictions at regional levels: Predictions regional level

Thank you for your attention!

carlotta.montorsi@liser.lu

Example clusters

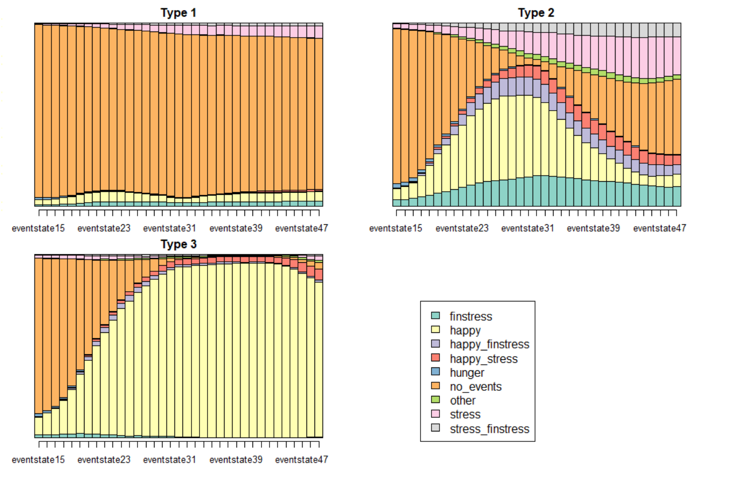
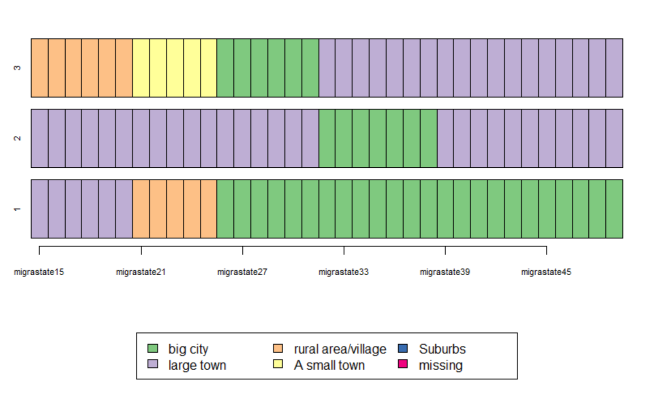


Figure 7: Clusters of housing arrangement, pooled sample

ID	age	Emotion: Type 1	Emotion: Type 2	Emotion: Type 3	
1	56	1	0	0	...
2	53	0	1	0	...
3	63	1	0	0	...
...	...	...	...	...	

sequence representation

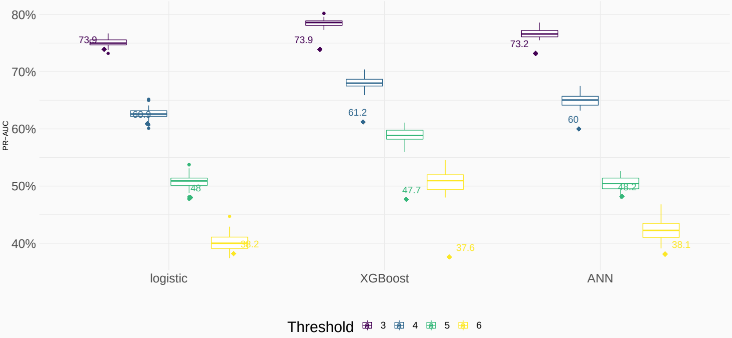
Example Unstructured



ID	Age15: Big city	Age15: Large Town	Age15: Small tows	Age15: Rural Area	Age15: Suburbs	Age15: Missing	...
1	0	1	0	0	0	0	...
2	0	1	0	0	0	0	...
3	0	0	0	1	0	0	...
...	...	...	...	...	...	...	...

sequence representation

Sensitivity across sub-samples



Conclusion

Predictions: European Regions

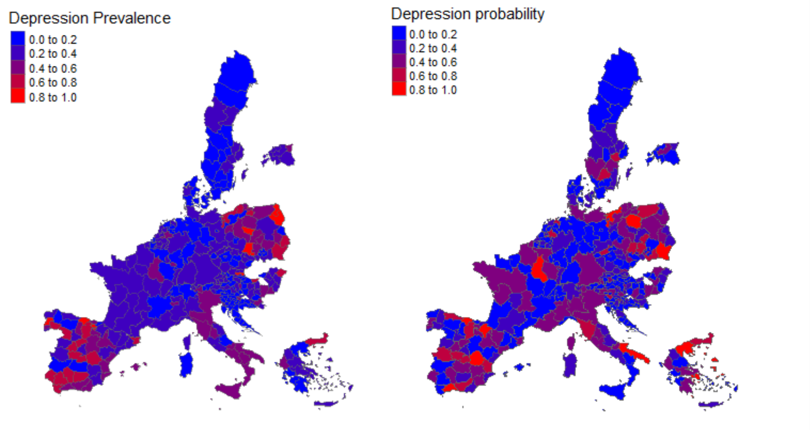


Figure 8: Left: observed depression rate at NUTS3 level. Right: aggregated depression probabilities at NUTS3

Conclusion