Predicting Depression in Old Age: combining life course data with machine learning

Well-Being Conference

June 1-4, 2022, Luxembourg

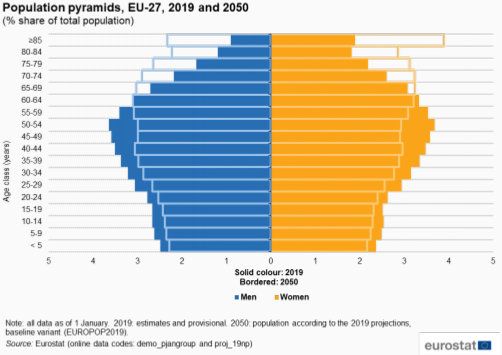
Carlotta Montorsi

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 is an independent predictor of other major diseases: Alzheimer, dementia and diabetes
- Depression is costly. Annual cost of depression in Europe: 253 euros per inhabitants → 1% of the total economy of Europe (P. Sobocki, B. Jönsson, J. Angst, C. Rehnberg, 2006)
- Depression in old age is both under-diagnosed and under-treated in primary care settings

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
    - 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
  - Lack of bio-markers/risk factors
  - Humans subjectivity
- Could we predict clinical depression from past life trajectories? Which data do we need?

Objectives

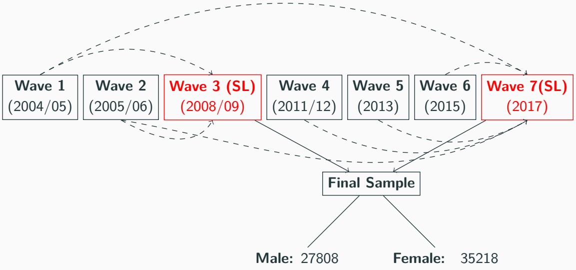
- Prevention strategies and improvements in early identification are essential (WHO, 2016).
- Predicting depression is a challenge
  - Lack of bio-markers/risk factors
  - Humans subjectivity
- Could we predict clinical depression from past life trajectories? Which data do we need?
- Are there differences in depression patterns among females and males?

Table of contents

Data

Data Source/Subjects

- The Survey of Health, Ageing and Retirement in Europe (SHARE).
- **Retrospective information** are collected in SHARELIFE (SL) questionnaire: wave 3 and wave 7.



- We discard:
    1. respondents aged 89 +. Problem of recall bias.
    2. respondents that provide little attention during the interview
    3. respondents with missing variables in depression symptoms across all waves

Measurements framework

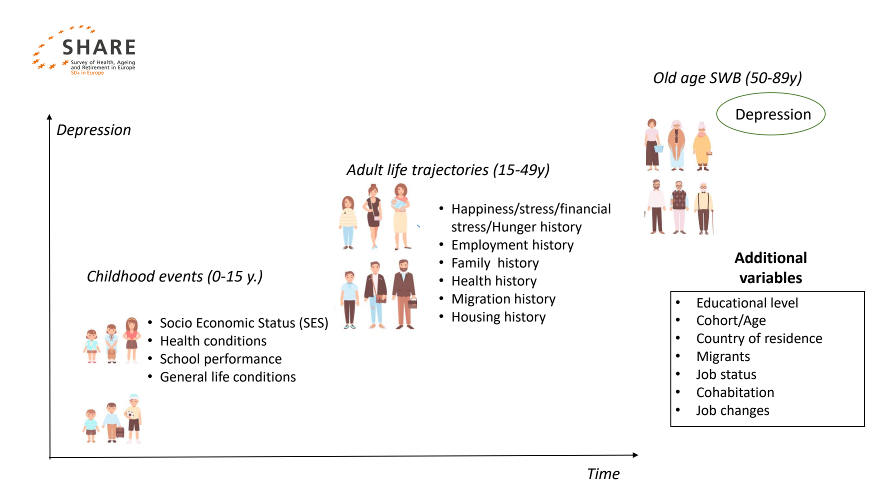


Figure 1: Measurements framework

Depression in SHARE

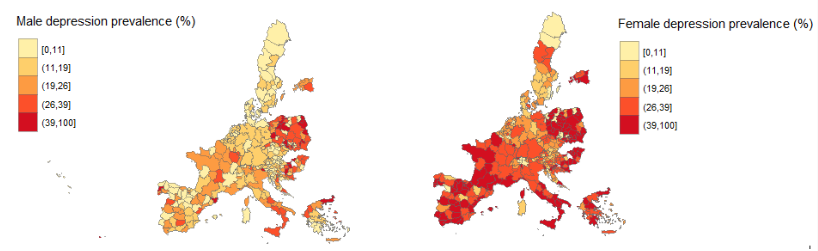


Figure 2: Depression prevalence across genders. Colors represent ventiles of the depression distributions in the pooled sample

- 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 4 or higher is categorized as case of depression (1) and a scale score below 4 as not depressed (0) (M. Prince et al., 1999b; E. Castro-Costa, M. Dewey et al.,2008)

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 .
- Two approaches to represent a life trajectory:
  1. Indicators of trajectories, such as whether the respondent experienced the birth of a child and when this event occurred
     - interpretation of the results (+)
     - limited number (-)
     - reduction of complexity (-)

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 .
- Two approaches to represent a life trajectory:
    1. Indicators of trajectories, such as whether the respondent experienced the birth of a child and when this event occurred
        - interpretation of the results (+)
        - limited number (-)
        - reduction of complexity (-)
    2. **Sequences** (A. Abbot, 1995):
        2.1 Holistic view over the life course (+)
        2.2 More information (+)
        2.3 Suitable for different representations (+)
        2.4 Less interpretable than typical predictors (-)

Sequences

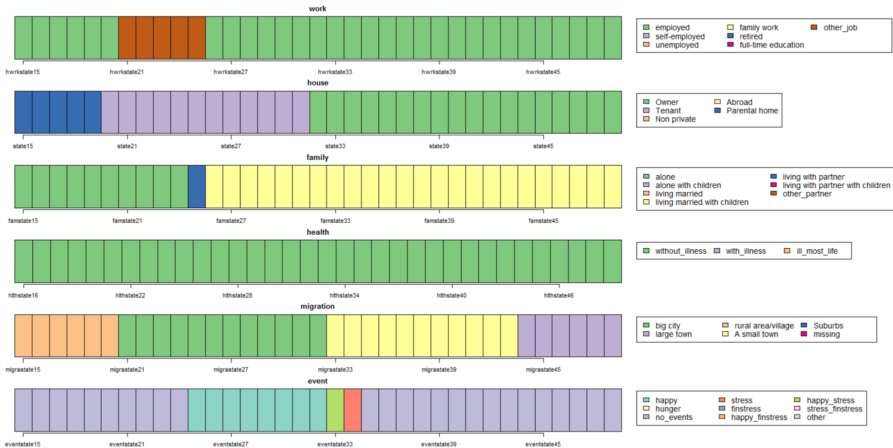


Figure 3: Individual sequences for the six analysed variables

Sequences representation

- In Social Science applications, sequences have been made operational in two ways:
  1. Clusters or Type: distinct groups of individuals' having similar life patterns Example cluster
  2. Sequences features: timing, duration, sequencing, entropy (M. Studer, G. Ritschard, 2016, D. Bolano et al. 2020) Example features
- We try also an unstructured representation. Example unstructured

Methods

Machine learning methods

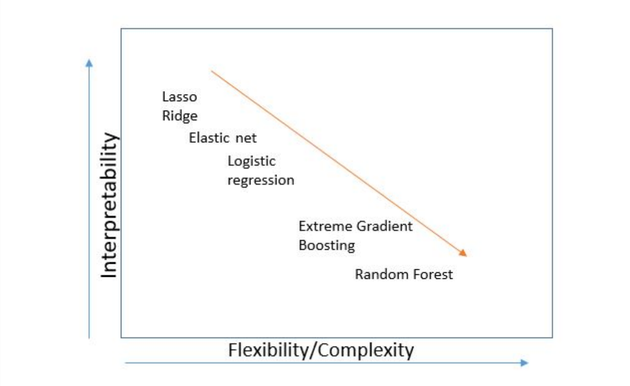


Figure 4: Predictive models explored in the analysis

Optimization (1)

Train-test split approach:

1. Training sets: 80% sample; test set: 20% sample

Optimization (1)

Train-test split approach:

1. Training sets: 80% sample; test set: 20% sample
2. Models' hyper-parameters: random/grid search + stratified 10-folds cross validation

Optimization (1)

Train-test split approach:

1. Training sets: 80% sample; test set: 20% sample
2. Models' hyper-parameters: random/grid search + stratified 10-folds cross validation
3. Select models hyper-parameters that maximize the Area Under the ROC curve

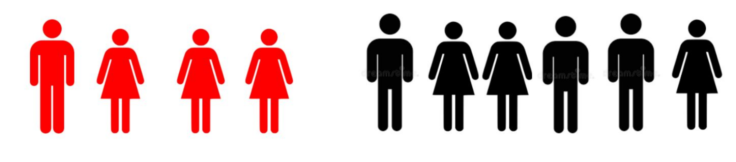
Optimization (1)

Train-test split approach:

1. Training sets: 80% sample; test set: 20% sample
2. Models' hyper-parameters: random/grid search + stratified 10-folds cross validation
3. Select models hyper-parameters that maximize the Area Under the ROC curve
4. Compare models' performance on the test set: sensitivity and accuracy

Predictive Performance

Population: depressed, **not depressed**

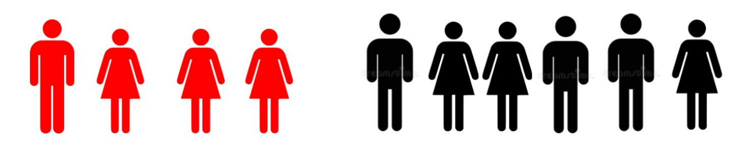


Prediction: depressed, **not depressed**

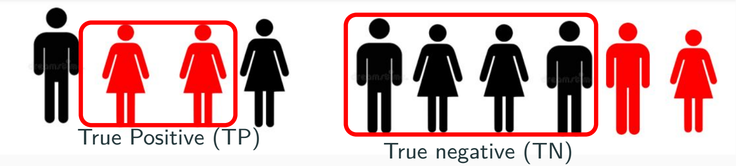
- Sensitivity: $\frac{TP}{TP+FN} \rightarrow \frac{2}{2+2} = 0.5$
- Accuracy: $\frac{TP+TN}{TP+TN+FP+FN} \rightarrow \frac{2+4}{10} = 0.6$

Predictive Performance

Population: depressed, **not depressed**

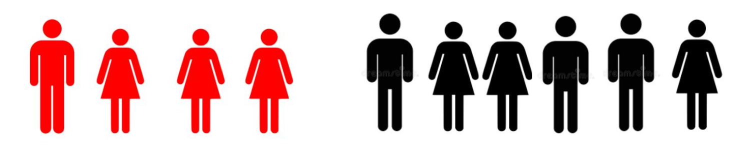


Prediction: depressed, **not depressed**



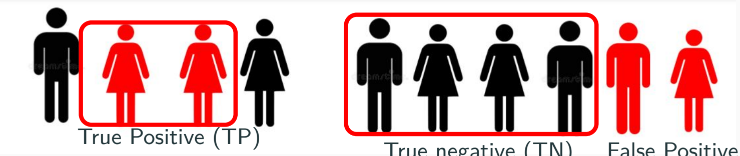
- Sensitivity: $\frac{TP}{TP+FN} \rightarrow \frac{2}{2+2} = 0.5$
- Accuracy: $\frac{TP+TN}{TP+TN+FP+FN} \rightarrow \frac{2+4}{10} = 0.6$

Population: depressed, **not depressed**



Prediction: depressed, **not depressed**

False Negative (FN)



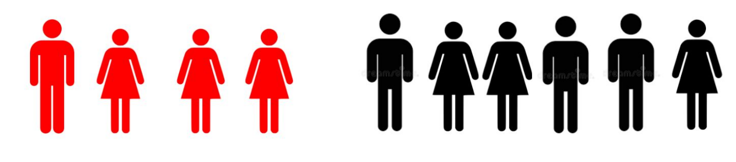
True Positive (TP)

True negative (TN)

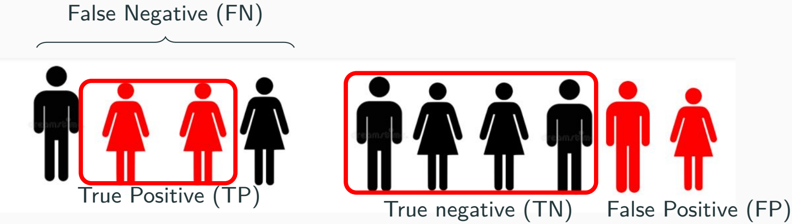
False Positive (FP)

- Sensitivity: $\frac{TP}{TP+FN} \rightarrow \frac{2}{2+2} = 0.5$
- Accuracy: $\frac{TP+TN}{TP+TN+FP+FN} \rightarrow \frac{2+4}{10} = 0.6$

Population: depressed, **not depressed**



Prediction: depressed, **not depressed**



- Sensitivity: $\frac{TP}{TP+FN} \rightarrow \frac{2}{2+2} = 0.5$
- Accuracy: $\frac{TP+TN}{TP+TN+FP+FN} \rightarrow \frac{2+4}{10} = 0.6$

Results

Models' sensitivity

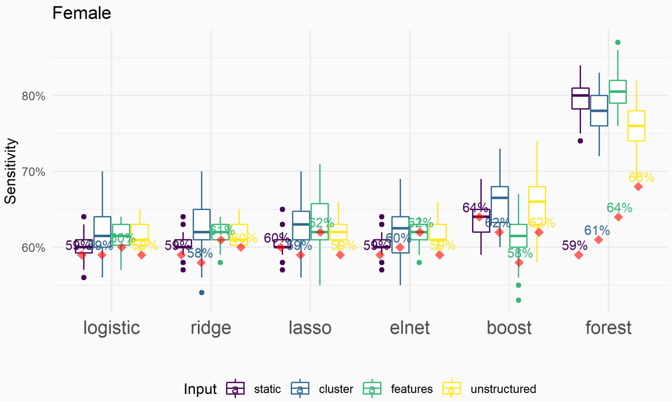


Figure 5: Sensitivity across models and input structures. Red dots indicates TEST error, box plots the distribution of 10-folds training errors. Female sample

- Sensitivity of the random forest increases along with the increasing dimensionality of the input structure.
- The random forest combined with the unstructured sequence representation achieves 68% of sensitivity in the test sample

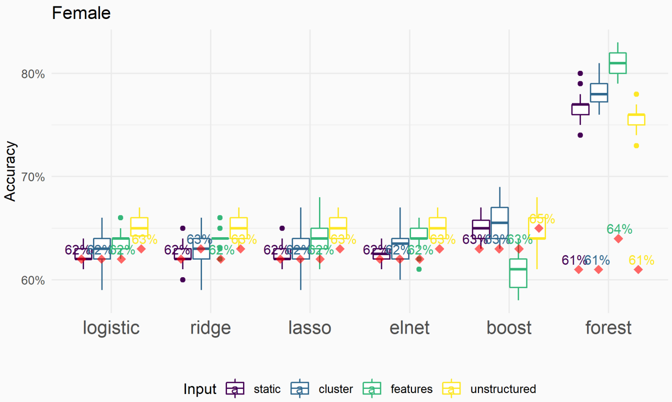


Figure 6: Sensitivity across models and input structures. Red dots indicates TEST error, box plots the distribution of 10-folds validation errors across countries. Male sample

- The random forest combined with the sequence features achieve the highest test sensitivity and accuracy.
- Random forest: problem of overfitting

Shapley values

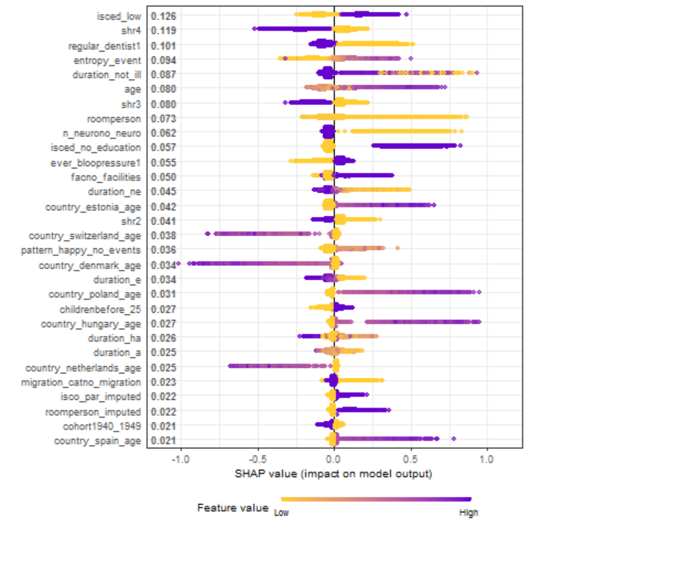


Figure 7: Top 30. Random forest Shapely values, female sample

Shapley values

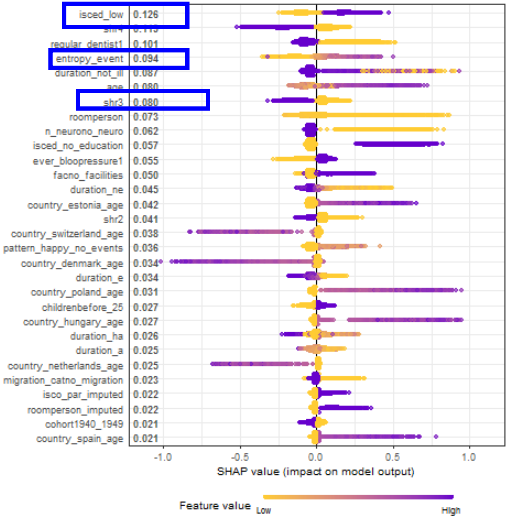


Figure 7: Top 30. Random forest Shapely values, female sample

Shapley values

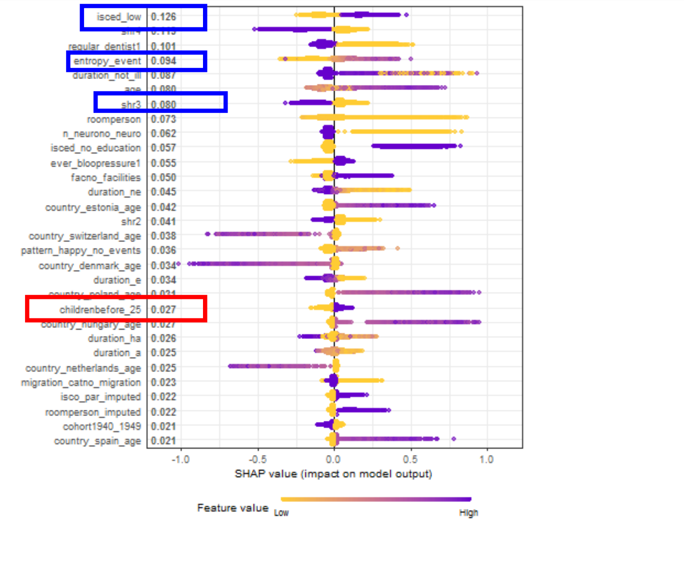


Figure 7: Top 30. Random forest Shapely values, female sample

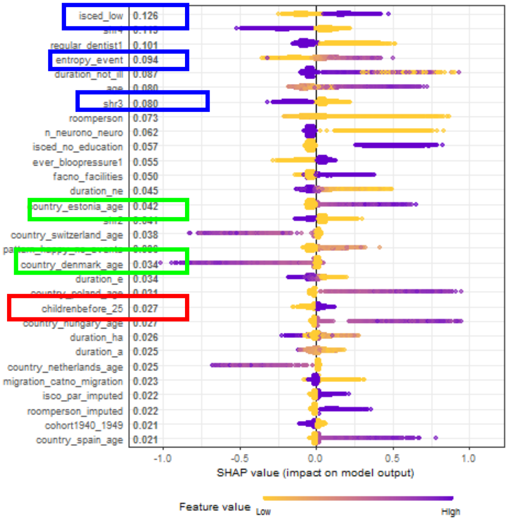


Figure 7: Top 30. Random forest Shapely values, female sample

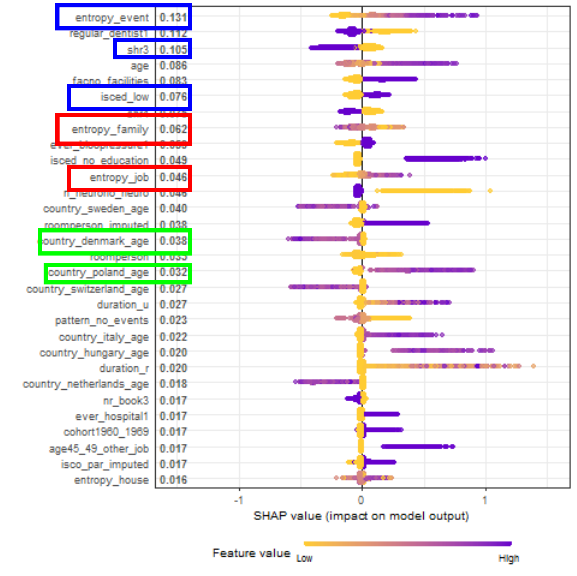


Figure 8: Top 30. Random forest Shapely values, male sample

Conclusion

- Depression is predictable from past life trajectories, up to a certain threshold
- The data required for achieving the highest predictive performance is more complex than what has been traditionally used in well-being studies
- We identify idiosyncratic and common patterns across genders
- Interpretable machine learning tools may support the hypothesis creation process

Thank you for your attention!

carlotta.montorsi@liser.lu

Example clusters

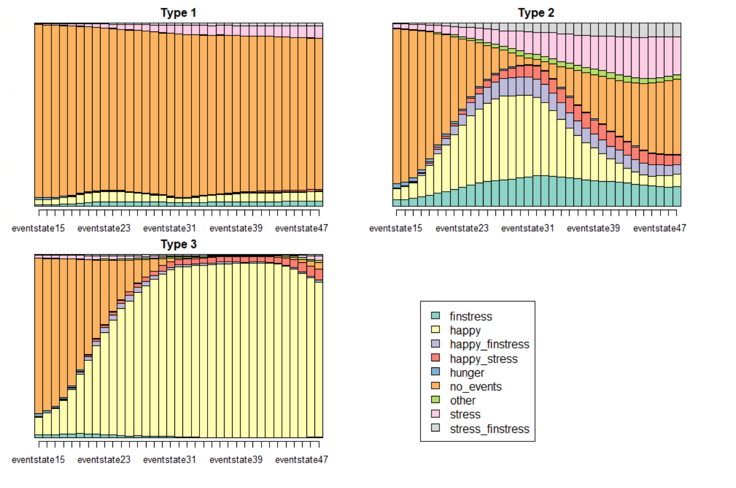
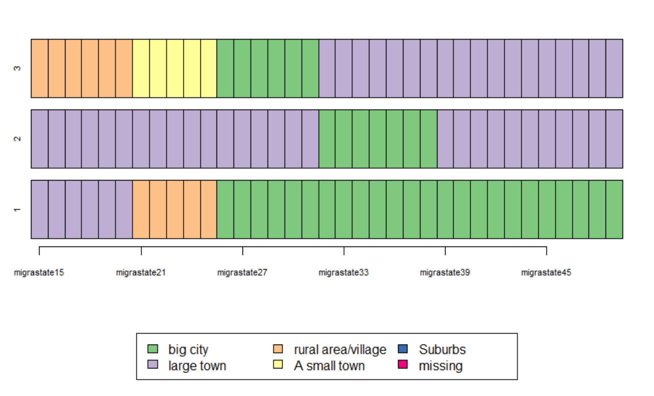


Figure 9: Clusters of emotional events, female sample

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

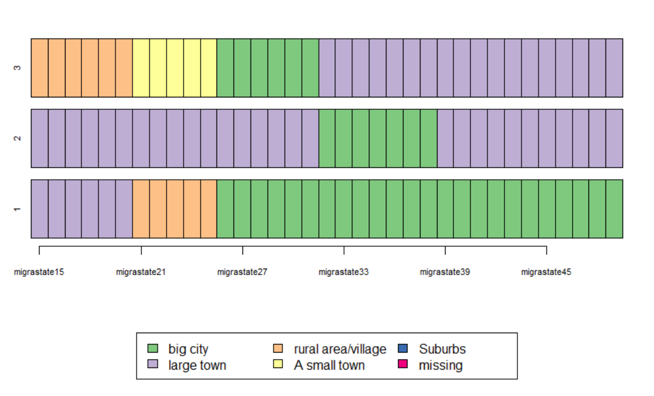
Example Features



ID	Duration BC	Duration ST	Duration Rur	LT → BC	LT → Rur → BC	Age(20-25) Rur	Entropy	
1	24	0	5	1	1	1	0.76	...
2	7	1	0	1	0	0	0.5	...
3	6	5	6	0	0	0	1.25	...
...	...	...	...	...	...	...	...	...

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