A MOBILE APP FOR COMPARATIVE FITNESS EVALUATION

Lena Zess, Christian Beyer, Myra Spiliopoulou
Otto von Guericke University Magdeburg (DE)

AIA Forschungsverbund Autonomie im Alter
EUROPÄISCHE UNION EFRE Europäischer Fonds für regionale Entwicklung
SACHSEN-ANHALT

OTTO VON GUERICKE UNIVERSITÄT MAGDEBURG
INF FAKULTÄT FÜR INFORMATIK
MEDIZINISCHE FAKULTÄT

Abstract

The aim of this project was to create an Android app that asks the user several health questions and then evaluates his/her age-based relative fitness. Therefore it uses a cluster model that was learned from a health survey data set. The app also visually juxtaposes the cluster properties and the user‘s answers to help him/her understand the results.

The health data set

- Questionnaire data set from ‘The 10,000 Immunomes Project‘ [1]
  - 1,422 participants
  - Questions: age, gender, race, 54 health attributes → too many for K means clustering
- Extracted 5 binary attributes
  - Abnormal blood cell count (CSQ1-4)*
  - Severe pulmonary disease (CSQ11-13)*
  - Treatment for cancer in the past 5 years
  - Arthritis or rheumatism for more than 3 months
  - Weakness for more than 3 months (CSQ41-43)*

  ** combines multiple attributes using inclusive disjunction*

… plus the age attribute (discretised)

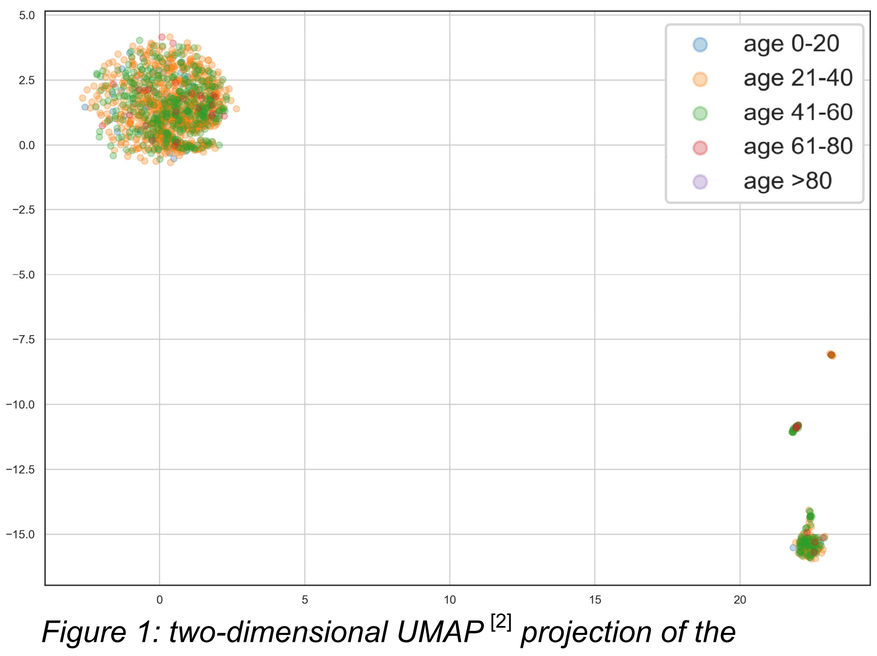

Figure 1: two-dimensional UMAP [2] projection of the five-dimensional data set coloured by age groups

- Structure analysis: Two-dimensional projection of the five-dimensional data set shows 4 clear groups with different age composition
- Only the health-related attributes are used for the following clustering. No clustering by age.

Clustering and Finding the optimal K

- Built several cluster models using Weka [3]
  - Clustering method: SimpleKMeans
  - Distance metric: Euclidean distance
  - Number of clusters K: 2, 3, ..., 10, 11
- Elbow point method: compared the within cluster SSE of the different models to find the optimal K → chose K = 4: is elbow point and matches structure from figure 1 best

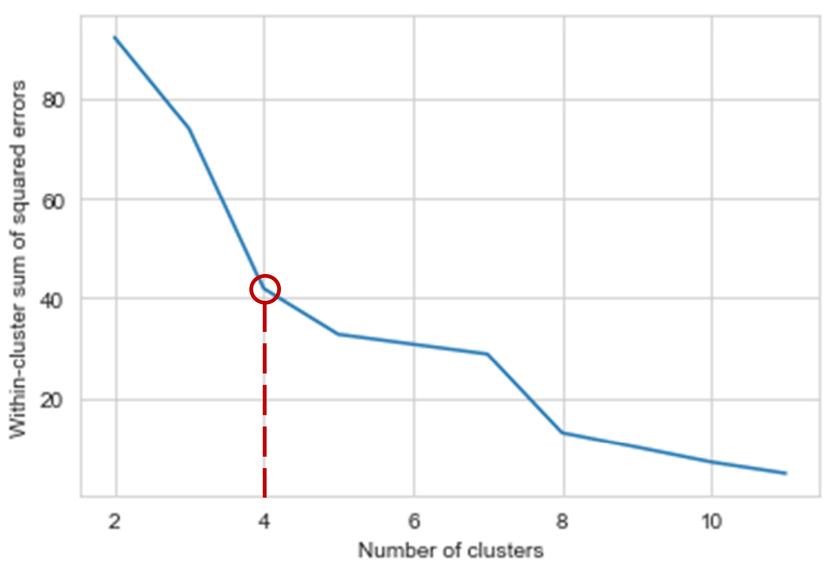

Figure 2: within cluster SSE in relation to the number of clusters

Figure 3: screenshots of the app showing questionnaire (top) and evaluation (bottom)

The application

- For Android smartphones or tablets
- User can input age and if he/she has some of the 5 medical conditions
- Uses Weka to assign the user to one of the clusters based on his/her answers
- Evaluates fitness:
  - Here fitness is defined as the difference between the user‘s age and the average age group of his/her cluster. For example if the user is 30 years old but belongs to a cluster with an average age of 41-60, he/she is in relatively poor health.
  - Depending on the result of this comparison the app displays a message with congratulations, a neutral or a concerned comment.
- Visualises the user‘s answers next to the cluster properties to make the cluster assignment comprehensible (see below)

The cluster-user diagram

- One example for this diagram is shown in the second screenshot of figure 3. The diagram will adapt to the user’s answers.
- X axis:
  - First 4 columns stand for the clusters
  - Last column stands for the user’s answers
- Y axis: The pictograms represent the medical conditions (same order as in the list on the left).
- Circles: Size and colour both express the probability that we observe the respective medical condition on an individual within the respective cluster. If there is no visible circle this represents a 0% probability. The right column with the user’s answers will always show a mix of either big yellow (100%) or non-visible (0%) circles as the answers are definite.

Acknowledgements

To Prof. Dr. Monika Brunner-Weinzierl from the medical faculty of Otto von Guericke University Magdeburg for the cooperation in the context of the EFRE-funded ImmunLearning project

References

[1] Zalocusky KA, Kan MJ, Hu Z, Dunn P, Thomson E, Wiser J, Bhattacharya S, Butte AJ: The 10,000 Immunomes Project: Building a Resource for Human Immunology. Cell reports 25(2), 513-522 (2018)

[2] McInnes L, Healy J: Uniform Manifold Approximation and Projection for Dimension Reduction (UMAP), v0.5. Available: https://umap-learn.readthedocs.io/en/latest/, Accessed on: Apr 28, 2021

[3] Hall M: Waikato Environment for Knowledge Analysis (Weka), v3.8.4. Available: https://www.cs.waikato.ac.nz/~ml/weka/, Accessed on: Apr 28, 2021