Intelligent Assistance for Expert-Driven Subpopulation Discovery in High-Dimensional Time-Stamped Medical Data

Uli Niemann

22.06.2020

Contents

Structure

1 Thesis Overview
1.1 Motivation
1.2 Medical Applications
1.3 Open Challenges
1.4 Structure and Summary of Contributions

PART I SUBPOPULATION DISCOVERY

2 Interactive discovery and inspection of subpopulations
2.1 The SHIP dataset
2.2 Motivation and comparison with related work
2.3 Subpopulation discovery workflow and interactive mining assistant
2.4 Experiments and findings
2.5 Benefits of our workflow

3 Identifying diverse subpopulations
3.1 Motivation and comparison with related work
3.2 Finding diverse classification rules
3.3 Experiments and findings
3.4 Benefits our our method

PART II EXPLOITING TEMPORAL INFORMATION

4 Constructing evolution features to capture change over time
4.1 Motivation and comparison with related work
4.2 Overview of the mining workflow
4.3 Generating evolution features
4.4 Experiments and findings
4.5 Benefits our our method

5 Feature extraction from short temporal sequences for clustering
5.1 The diabatic foot dataset
5.1 Motivation and comparison with related work 5.2 Overview of the mining workflow
5.3 Experiments and findings
5.4 Benefits our our method

PART III POST-MINING FOR INTERPRETATION

Chapter 1

Thesis overview

1.1 Motivation

1.2 Medical applications

1.3 Open challenges

1.4 Structure and summary of contributions

Part I

SUBPOPULATION DISCOVERY IN HIGH-DIMENSIONAL DATA

Chapter 2

Interactive discovery and inspection of subpopulations

Based on: [1]

2.1 Motivation and comparison with related work

2.2 The SHIP dataset

2.3 Subpopulation discovery workflow and interactive mining assistant

2.4 Experiments and findings

2.5 Benefits of our workflow

Chapter 3

Identifying diverse subpopulations

Based on: [2]

3.1 Motivation and comparison with related work

3.2 Finding diverse classification rules

3.3 Experiments and findings

3.4 Benefits our our method

Part II

EXPLOITING TEMPORAL INFORMATION

Chapter 4

Constructing evolution features to capture change over time

Based on: [3]

4.1 Motivation and comparison with related work

4.2 Overview of the mining workflow

4.3 Generating evolution features

4.4 Experiments and findings

4.5 Benefits our our method

Chapter 5

Feature extraction from short temporal sequences for clustering

5.1 Motivation and comparison with related work

Based on: [4]–[6]

5.2 The diabatic foot dataset

5.3 Overview of the mining workflow

5.4 Experiments and findings

5.5 Benefits our our method

Part III

POST-MINING FOR INTERPRETATION

Chapter 6

Post-hoc interpretation of classification models

Based on: [7]–[10]

6.1 Motivation and comparison with related work

6.2 The aneurysm dataset

6.3 The tinnitus dataset

6.4 Overview of the mining workflow

6.5 Experiments and findings on aneurysm data

6.6 Experiments and findings on tinnitus data

6.7 Benefits our our method

Part IV

SUMMARY

Chapter 7

Conclusion and future work

7.1 Research results for medical expert-guided knowledge discovery

7.2 Future work

Part V

APPENDIX

Bibliography

[1] U. Niemann, H. Völzke, J.-P. Kühn, and M. Spiliopoulou, "Learning and inspecting classification rules from longitudinal epidemiological data to identify predictive features on hepatic steatosis," *Expert Systems with Applications*, vol. 41, no. 11, pp. 5405–5415, 2014, ISSN: 0957-4174. DOI: 10.1016/j.eswa.2014.02.040. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S095741741400116X.

[2] U. Niemann, M. Spiliopoulou, B. Preim, T. Ittermann, and H. Völzke, "Combining Subgroup Discovery and Clustering to Identify Diverse Subpopulations in Cohort Study Data," in *Proc. of IEEE Int. Symposium on Computer-Based Medical Systems (CBMS)*, 2017, pp. 582–587. DOI: 10.1109/CBMS.2017.15.

[3] U. Niemann, T. Hielscher, M. Spiliopoulou, H. Völzke, and J.-P. Kühn, "Can we classify the participants of a longitudinal epidemiological study from their previous evolution?" In *Proc. of IEEE Int. Symposium on Computer-Based Medical Systems (CBMS)*, 2015, pp. 121–126. DOI: 10.1109/CBMS.2015.12.

[4] U. Niemann, M. Spiliopoulou, F. Samland, T. Szczepanski, J. Grützner, A. Ming, J. Kellersmann, J. Malanowski, S. Klose, and P. R. Mertens, "Learning Pressure Patterns for Patients with Diabetic Foot Syndrome," in *Proc. of IEEE Int. Symposium on Computer-Based Medical Systems (CBMS)*, 2016, pp. 54–59. DOI: 10.1109/CBMS.2016.31.

[5] U. Niemann, M. Spiliopoulou, T. Szczepanski, F. Samland, J. Grützner, D. Senk, A. Ming, J. Kellersmann, J. Malanowski, S. Klose, and P. R. Mertens, "Comparative Clustering of Plantar Pressure Distributions in Diabetics with Polyneuropathy May Be Applied to Reveal Inappropriate Biomechanical Stress," *PLOS ONE*, vol. 11, no. 8, pp. 1–12, 2016. DOI: 10.1371/journal.pone.0161326. [Online]. Available: http://dx.doi.org/10.1371%2Fjournal.pone.0161326.

[6] U. Niemann, M. Spiliopoulou, J. Malanowski, J. Kellersmann, T. Szczepanski, S. Klose, E. Dedonaki, I. Walter, A. Ming, and P. R. Mertens, "Plantar temperatures in stance position: A comparative study with healthy volunteers and diabetes patients diagnosed with sensoric

neuropathy," *EBioMedicine*, vol. 54, p. 102 712, 2020, ISSN: 2352-3964. DOI: https://doi.org/10.1016/j.ebiom.2020.102712. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S2352396420300876.

[7] U. Niemann, P. Berg, A. Niemann, O. Beuing, B. Preim, M. Spiliopoulou, and S. Saalfeld, "Rupture Status Classification of Intracranial Aneurysms Using Morphological Parameters," in *Proc. of IEEE Int. Symposium on Computer-Based Medical Systems (CBMS)*, 2018, pp. 48–53. DOI: 10.1109/CBMS.2018.00016.

[8] U. Niemann, P. Brueggemann, B. Boecking, B. Mazurek, and M. Spiliopoulou, "Development and internal validation of a depression severity prediction model for tinnitus patients based on questionnaire responses and socio-demographics," *Scientific Reports*, vol. 10, no. 1, p. 4664, 2020, ISSN: 2045-2322. DOI: 10.1038/s41598-020-61593-z. [Online]. Available: https://doi.org/10.1038/s41598-020-61593-z.

[9] U. Niemann, B. Boecking, P. Brueggemann, W. Mebus, B. Mazurek, and M. Spiliopoulou, "Tinnitus-related distress after multimodal treatment can be characterized using a key subset of baseline variables," *PLOS ONE*, vol. 15, no. 1, pp. 1–18, Jan. 2020. DOI: 10.1371/journal.pone.0228037. [Online]. Available: https://doi.org/10.1371/journal.pone.0228037.

[10] U. Niemann, B. Boecking, P. Brueggemann, B. Mazurek, and M. Spiliopoulou, "Gender-Specific Differences in Patients With Chronic Tinnitus—Baseline Characteristics and Treatment Effects," *Frontiers in Neuroscience*, vol. 14, p. 487, 2020, ISSN: 1662-453X. DOI: 10.3389/fnins.2020.00487. [Online]. Available: https://www.frontiersin.org/article/10.3389/fnins.2020.00487.