Table of Contents

Acknowledgements ix

Abstract 1

Resumen 3

Resum 5

1 Introduction 7
1.1 Motivation 7
1.2 Research questions and objectives 8
1.3 Thesis contributions 10
1.4 Projects and partners 14
1.5 Outline 15

2 Rationale 17
2.1 Data Quality. Multisource and Temporal Variability Assessment Methods 18
2.2 Machine Learning 22
2.3 Graphs 31

3 Journal article (i) 35
Pérez-Benito, F.J., Sáez, C., Conejero, J.A., Tortajada, S., Valdivieso, B., García-Gómez, J.M. (2019). Temporal variability analysis reveals biases in electronic health records due to hospital process reengineering interventions over seven years. *PLoS ONE*, 14(8): e0220369. 35
Abstract 35
3.1 Background and significance 36
3.2 Materials and Methods 37
3.3 Results 43
3.4 Discussion 49
3.5 Conclusions 54
3.6 Supplementary material 55

4 Journal article (ii) 63
Pérez-Benito, Conejero, J.A., Sáez, C., García-Gómez, J.M., Navarro-Pardo, E., Florencio, L.L., Fernández-de-las-Peñas, C. (2020). Subgrouping factors influencing migraine intensity in women: A semi-automatic methodology based on Machine Learning and Information Geometry. *Pain Practice*, 20(3) 297-309. 63
Abstract 63
4.1 Indroduction 64
4.2 Methods 66
4.3 Results 71
4.4 Discussion 75
4.5 Conclusion 78

5 Journal article (iii) 81
F.J. Pérez-Benito, F. Signol, J.C. Perez-Cortes, A. Fuster-Baggetto, M. Pollán, B. Pérez-Gómez, D. Salas-Trejo, M. Casals, I. Martínez, R. LLobet. (2020). A deep learning system to obtain the optimal parameters for a threshold-based breast and dense tissue segmentation. *Computer Methods and Programs in Biomededicine* 195, 105668 81
Abstract 81
5.1 Background 82
5.2 Methods 83
5.3 Results 91
5.4 Discussion 97
5.5 Conclusion 98

6 Journal article (iv) 101
Pérez-Benito, F.J., Villacampa-Fernández, P., Conejero, J.A., García-Gómez, J.M., Navarro-Pardo, E. (2019). A happiness degree predictor using the conceptual data structure for deep learning architectures. *Computer Methods and Programs in Biomedicine*, 168, 59-68. 101
Abstract 101
6.1 Introduction 102
6.2 Materials 105
6.3 Methods 108
6.4 Experimental Results 114
6.5 Discussion 118
6.6 Conclusions 121

7 Journal article (v) 123
Pérez-Benito, F.J., Conejero, J.A., García-Gómez, J.M., Navarro-Pardo, E. (2019). Community detection based deep neural network (CD-DNN) architectures: a fully automated framework based on Likert-scale data. *Mathematical Methods in the Applied Sciences*, 43, 8290-8301. 123
Abstract 123
7.1 Introduction 124
7.2 Proposed Methodology 125
7.3 Proposal of DNN architecture 128
7.4 Experimental Results 131
7.5 Discussion 136
7.6 Conclusion 137

8 Concluding remarks and recommendations 139
8.1 Concluding remarks 139
8.2 Recommendations 142

References 145