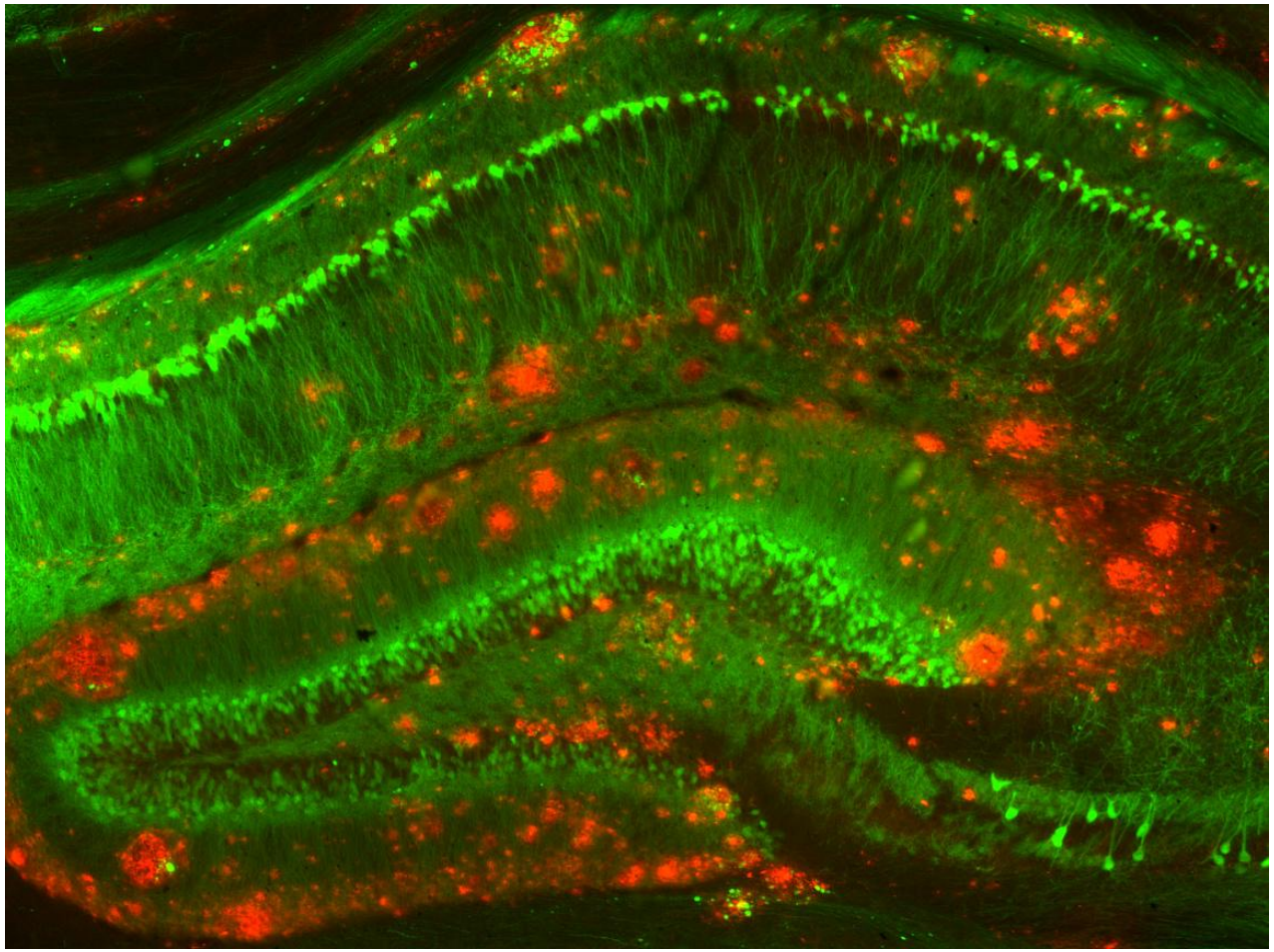
AmyloGram 2.0: MBO in the prediction of amyloid proteins

Dominik Rafacz

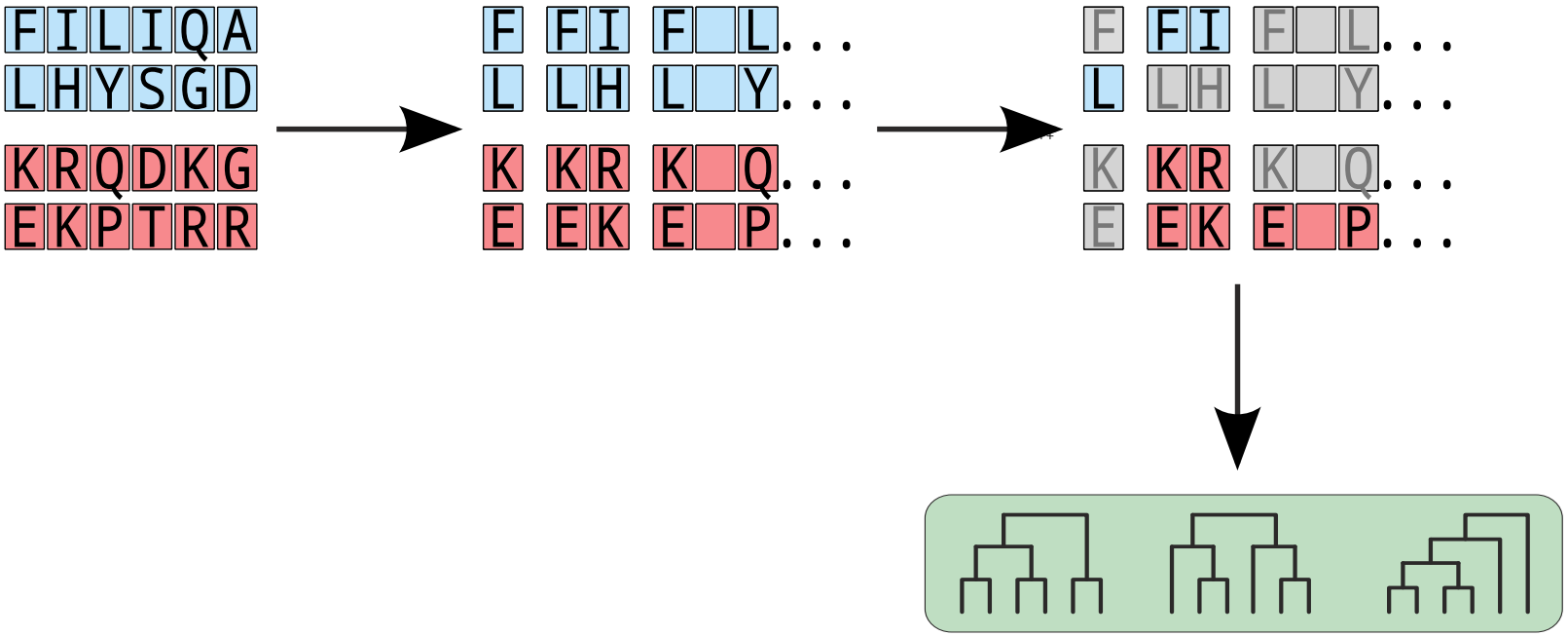
Warsaw University of Technology

Amyloidogenic proteins



Amyloid aggregates (red) around neurons (green). Strittmatter Laboratory, Yale University.

AmyloGram - n-grams analysis



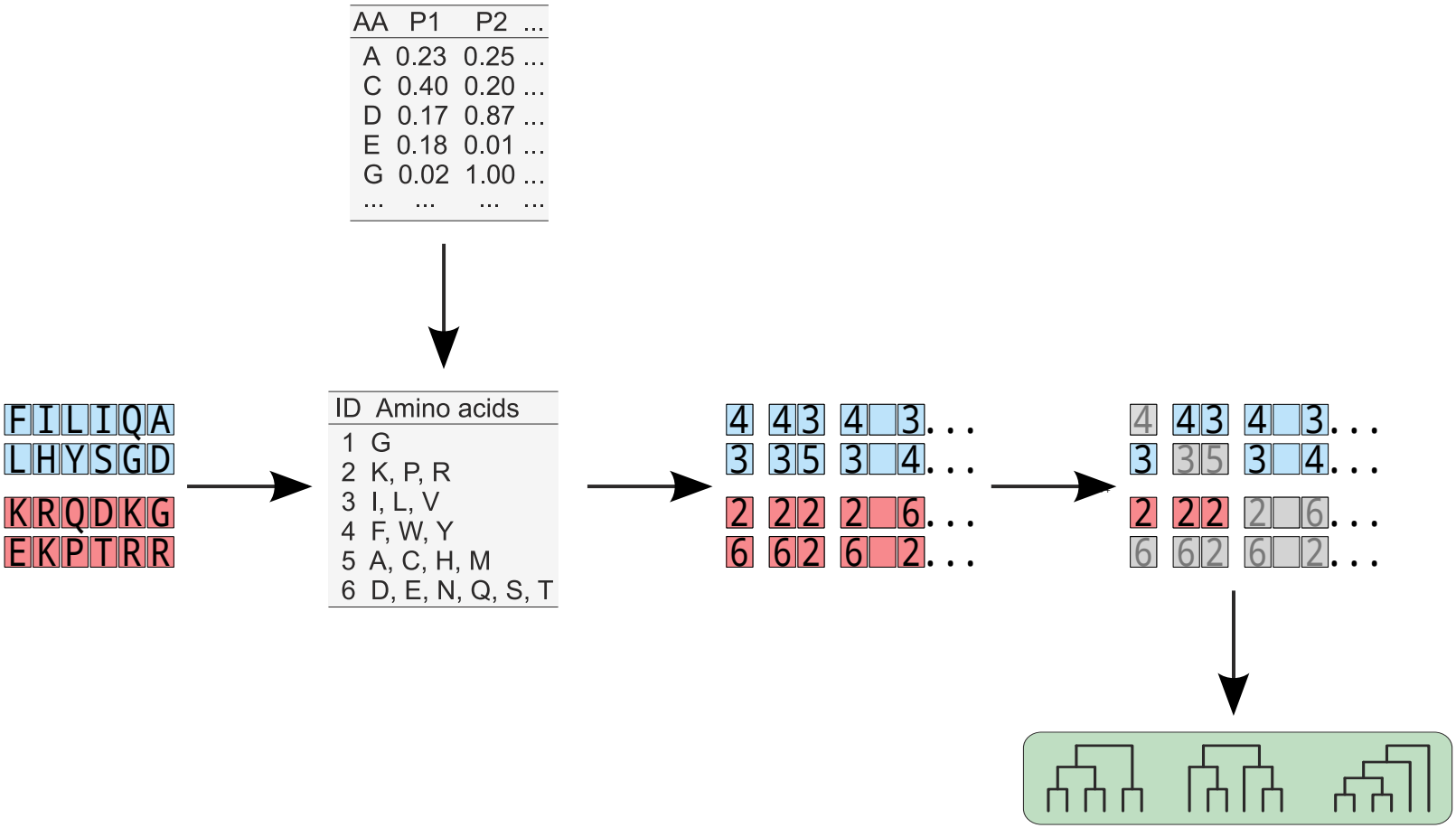
Example 1-grams: A, L, G

Example 2-grams: AL, MM, MY

Example 2-grams (with a gap): A-L, M-M, M-Y

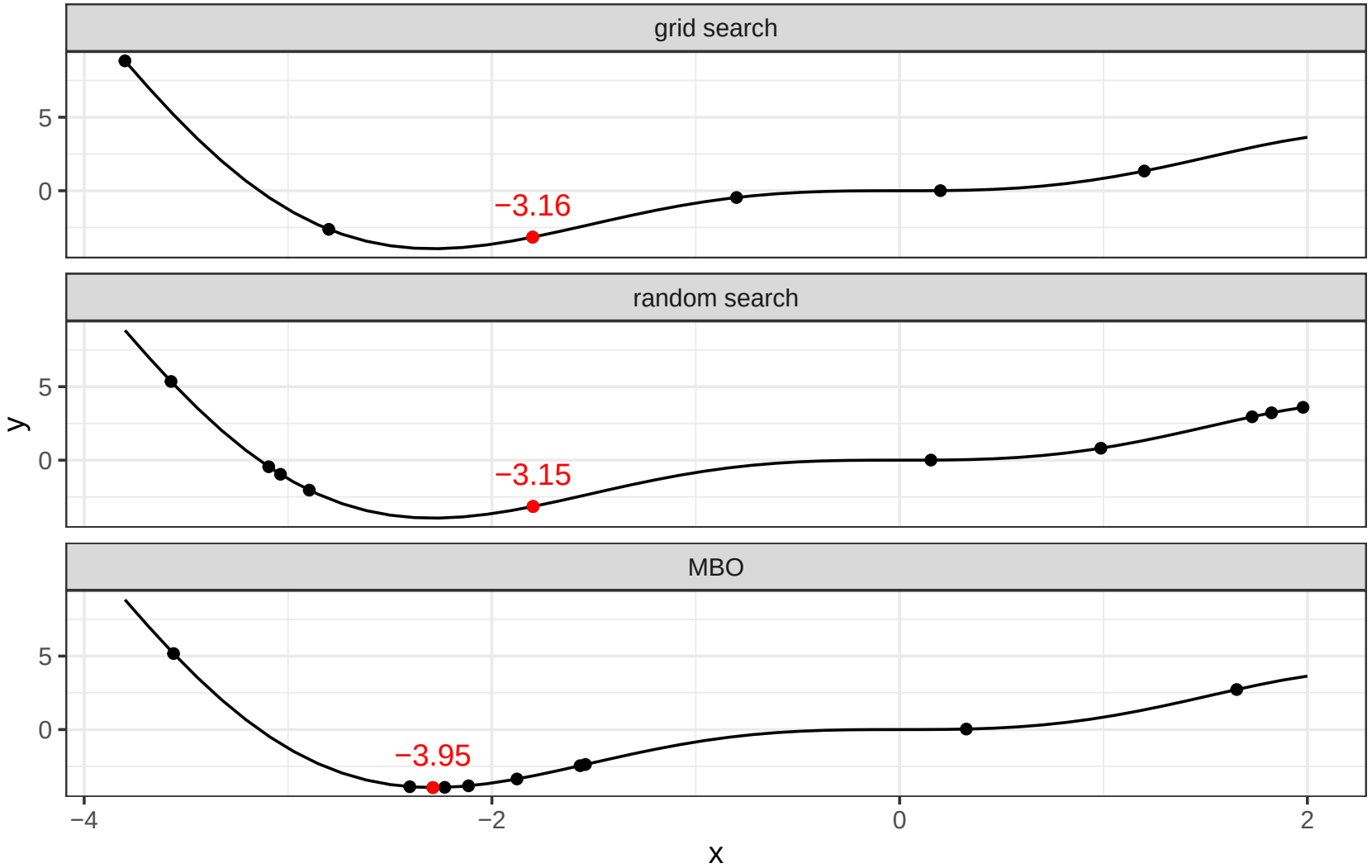
Burdukiewicz, M., Sobczyk, P., Rödiger, S., Duda-Madej, A., Mackiewicz, P., and Kotulska, M. (2017). Amyloidogenic motifs revealed by n-gram analysis. Scientific Reports 7, 12961

AmyloGram - alphabet reduction

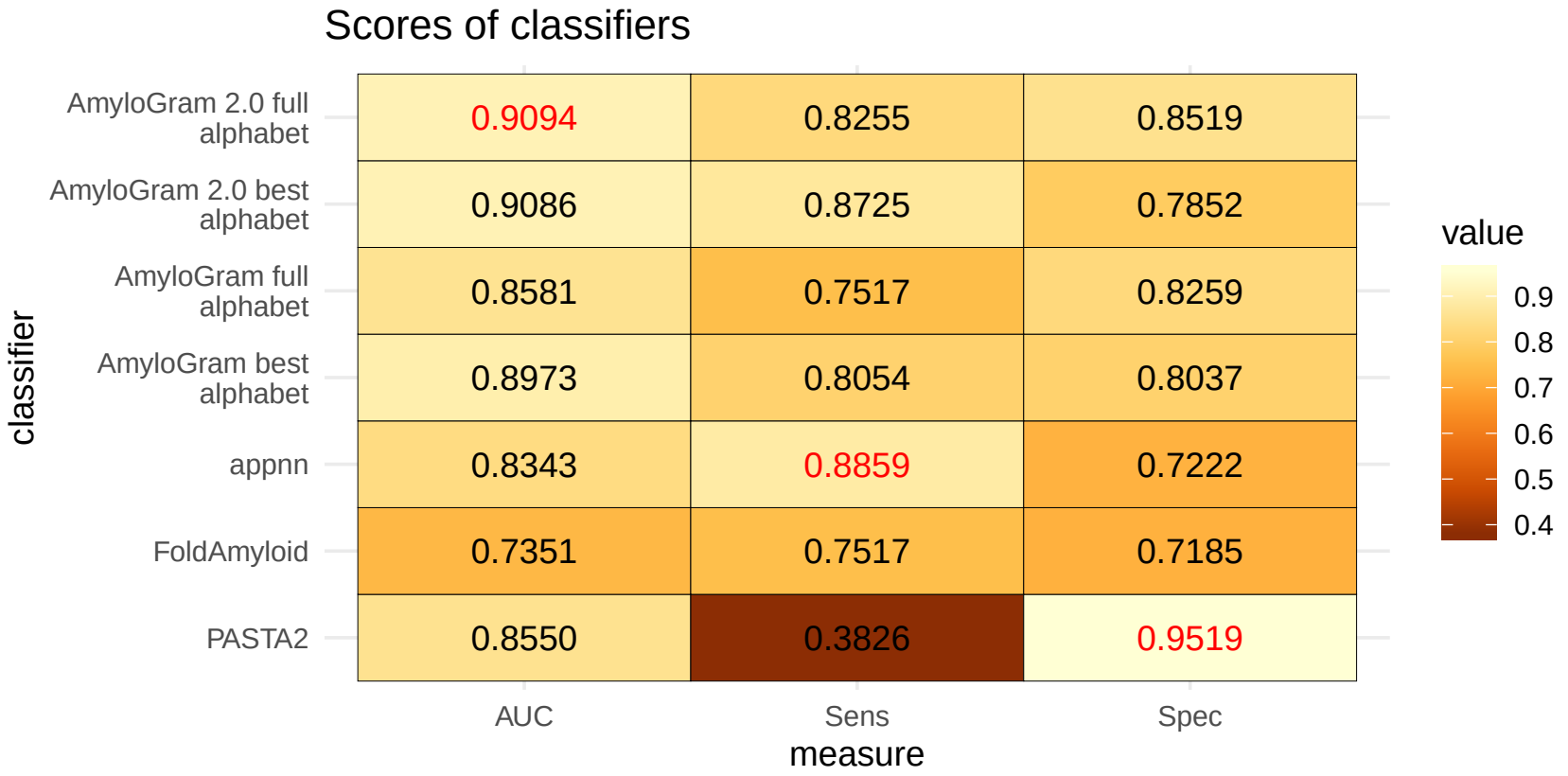


Burdukiewicz, M., Sobczyk, P., Rödiger, S., Duda-Madej, A., Mackiewicz, P., and Kotulska, M. (2017). Amyloidogenic motifs revealed by n-gram analysis. Scientific Reports 7, 12961

MBO



Results



Features used in the model in original model: **262** ()
Features used in the model with full alphabet: **33620** (100%)
Features used in the model with best alphabet: **1014**

Acknowledgements & references

Special thanks to:

- Students' Union of Warsaw University of Technology
- Michał Burdukiewicz (Warsaw University of Technology)

Burdukiewicz, M., Sobczyk, P., Rödiger, S., Duda-Madej, A., Mackiewicz, P., and Kotulska, M. (2017). Amyloidogenic motifs revealed by n-gram analysis. Scientific Reports 7, 12961

Bischl, B., Richter, J., Bossek, J., Horn, D., Thomas, J. and Lang, M. (2017). mlrMBO: A Modular Framework for Model-Based Optimization of Expensive Black-Box Functions.