[1] [2] [3, 4]

1. Strack B, DeShazo JP, Gennings C, Olmo JL, Ventura S, Cios KJ, Clore JN: **Impact of HbA1c measurement on hospital readmission rates: analysis of 70,000 clinical database patient records**. *Biomed Res Int* 2014, **2014**:781670.

2. Centers for Disease Control and Prevention: **National Diabetes Statistics Report, 2017**. In*.* Edited by Centers for Disease Control and Prevention US Department of Health and Human Services. Atlanta, GA; 2017.

3. Leppin AL, Gionfriddo MR, Kessler M, Brito JP, Mair FS, Gallacher K, Wang Z, Erwin PJ, Sylvester T, Boehmer K *et al*: **Preventing 30-day hospital readmissions: a systematic review and meta-analysis of randomized trials**. *JAMA Intern Med* 2014, **174**(7):1095-1107.

4. Rubin DJ, Handorf EA, Golden SH, Nelson DB, McDonnell ME, Zhao H: **Development and Validation of a Novel Tool to Predict Hospital Readmission Risk among Patients with Diabetes**. *Endocr Pract* 2016, **22**(10):1204-1215.

Note: the original paper (Strack) actually didn’t discuss their model fit at all (no PPV or AUC)

The Rubin paper (4) used a dataset from a single academic centre and their algorithm achieved an AUC of 0.69.

R-Package citations:

tidyverse

Hadley Wickham (2017). tidyverse: Easily Install and Load the 'Tidyverse'. R package version 1.2.1. <https://CRAN.R-project.org/package=tidyverse>

caret

Max Kuhn. Contributions from Jed Wing, Steve Weston, Andre Williams, Chris Keefer, Allan Engelhardt, Tony Cooper, Zachary Mayer, Brenton Kenkel, the R Core Team, Michael Benesty, Reynald Lescarbeau, Andrew Ziem, Luca Scrucca, Yuan Tang, Can Candan and Tyler Hunt. (2018). caret: Classification and Regression Training. R package version 6.0-81. <https://CRAN.R-project.org/package=caret>

xgboost

Tianqi Chen, Tong He, Michael Benesty, Vadim Khotilovich, Yuan Tang, Hyunsu Cho, Kailong Chen, Rory Mitchell, Ignacio Cano, Tianyi Zhou, Mu Li, Junyuan Xie, Min Lin, Yifeng Geng and Yutian Li (2018). xgboost: Extreme Gradient Boosting. R package version 0.71.2. <https://CRAN.R-project.org/package=xgboost>

skimr

Amelia McNamara, Eduardo Arino de la Rubia, Hao Zhu, Shannon Ellis and Michael Quinn (2018). skimr: Compact and Flexible Summaries of Data. R package version 1.0.3. <https://CRAN.R-project.org/package=skimr>

data.table

Matt Dowle and Arun Srinivasan (2018). data.table: Extension of `data.frame`. R package version 1.11.8. <https://CRAN.R-project.org/package=data.table>

mltools

Ben Gorman (2018). mltools: Machine Learning Tools. R package version 0.3.5. https://CRAN.R-project.org/package=mltools

corrplot

Taiyun Wei and Viliam Simko (2017). R package "corrplot": Visualization of a Correlation Matrix (Version 0.84). Available from <https://github.com/taiyun/corrplot>

ROCR

Sing T, Sander O, Beerenwinkel N, Lengauer T (2005). “ROCR: visualizing classifier performance in R.” *Bioinformatics*, **21**(20), 7881. <URL: <http://rocr.bioinf.mpi-sb.mpg.de>>.

DMwR

Torgo, L. (2010). Data Mining with R, learning with case studies Chapman and Hall/CRC. URL: <http://www.dcc.fc.up.pt/~ltorgo/DataMiningWithR>

t-SNE

Jesse H. Krijthe (2015). Rtsne: T-Distributed Stochastic Neighbor Embedding using a Barnes-Hut Implementation, URL: <https://github.com/jkrijthe/Rtsne>

glmnet

Jerome Friedman, Trevor Hastie, Robert Tibshirani (2010). Regularization Paths for Generalized Linear Models via Coordinate Descent. Journal of Statistical Software, 33(1), 1-22. URL <http://www.jstatsoft.org/v33/i01/>

tableone

Kazuki Yoshida and Justin Bohn. (2018). tableone: Create 'Table 1' to Describe Baseline Characteristics. R package version 0.9.3. <https://CRAN.R-project.org/package=tableone>

R

R Core Team (2018). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/