Prediction of Postpartum Depression Using Machine Learning methods

S. Ioannou[1], G.A. Holtman[2], L.L Peters[2,3], C.L.H. Bockting[4], H. Burger[2] , C.H. zu Eulenburg[1]

1. Department of Epidemiology, University Medical Center Groningen, University of Groningen, The Netherlands,
2. Department of General Practice & Elderly Care Medicine, University Medical Center Groningen, University of Groningen, The Netherlands,
3. Department of Midwifery Science AVAG, VU University Medical Center Amsterdam, Amsterdam Public Health Research Institute, The Netherlands,
4. Department of Psychiatry, Academic Medical Center, University of Amsterdam, The Netherlands

Introduction

Postpartum depression (PPD) is a clinical condition that appears in women even up to one year after childbirth. It has been recognized that PPD creates a significant impairment in functioning, that affects both the mother and the newborn and professional treatment is suggested.

Objectives

To apply and compare different machine learning methods in a practical clinical example: Predicting post-partum depression after 6 months of giving birth using data collected at onset of pregnancy (within twelve weeks).

Data and Methods

- Data was obtained from the Pregnancy and Depression (PAD) study[1].
- Data was partitioned into a training (70%) and test (30%) datasets. Models where created using the data from the training datasets and their predictive ability was tested on the test dataset.
- We tested the predictive ability of four methods, namely, logistic regression, decision tree, random forest and linear discriminant analysis.
- The models predictive ability is compared by their overall accuracy and ROC curves.

Results

The analyses included 2809 women, thereof 316 (13%%) with postpartum depression. The training set (test set) consists of 1966 (843) women, thereof with 210 (12%) presenting PPD in the training set and 106 (14%) in the test set.

Table 1. Odds ratios and 95% confidence intervals of LR model

	Outcome
	PPD as assessed by EPDS ≥10 at 6 months postnatally
Intercept	0.0004 (0.0000, 0.01)
STAI	1.01 (0.98, 1.03)
EPDS	1.16* (1.08, 1.25)
NEO - Neuroticism	1.09* (1.05, 1.13)
NEO - Extraversion	0.96* (0.92, 0.99)
NEO - Agreeableness	1.03 (0.99, 1.06)
NEO - Conscientiousness	1.07* (1.03, 1.11)
Married (Yes)	0.92 (0.64, 1.33)
Married (Missing)	0.31 (0.04, 5.92)
Paid Job (Yes)	0.86 (0.51, 1.49)
Paid Job (Missing)	1.14 (0.05, 11.84)
Higher Education (Yes)	0.94 (0.64, 1.39)
Higher Education (Missing)	2.60 (0.24, 19.86)
Comorbidity (Yes)	1.40 (0.85, 2.23)
Comorbidity (Missing)	1.19 (0.54, 2.46)
Current Psych. Treat. (Yes)	1.10 (0.56, 2.11)
Current Psych. Treat. (Missing)	1.26 (0.39, 3.38)
Feeling Supported (Yes)	0.84 (0.53, 1.37)
Feeling Supported (Missing)	0.0000 (-924.15, 225,823,227,088.00)
Observations	1,966
Log Likelihood	-522.22
Akaike Inf. Crit.	1,082.43

- The main predictors of the LR model are EPDS, neuroticism, extraversion and conscientiousness.

Figure 1. DT components on the test set

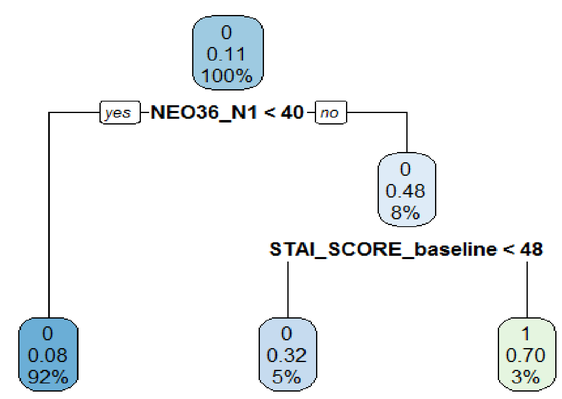


Figure 2. Variable importance table for RF

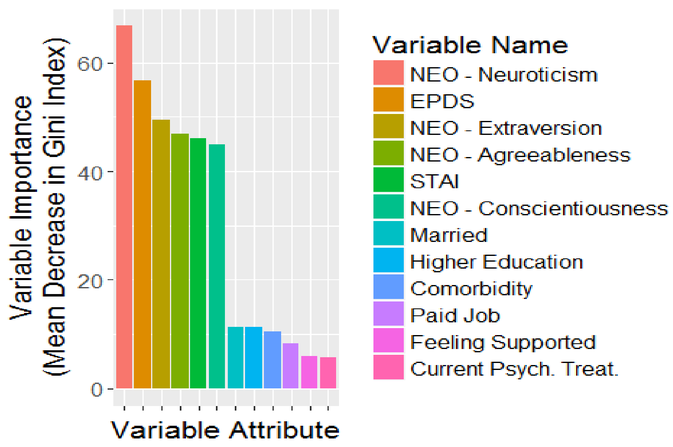
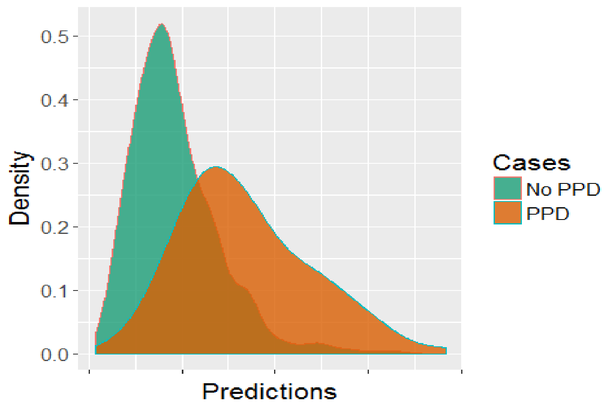


Figure 3. Stacked density plot of the discriminant of LDA



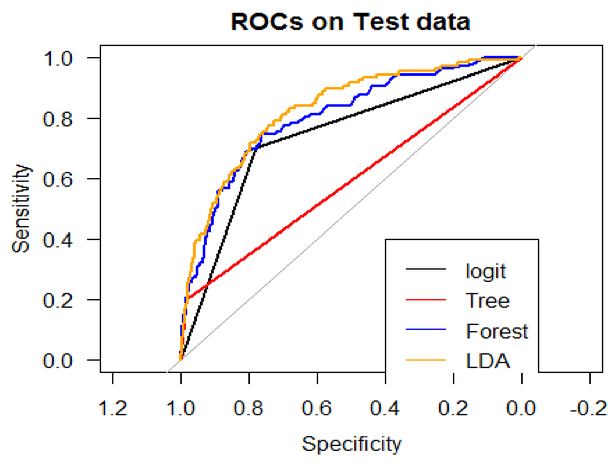
- DT partitions the data first based on neuroticism and then on the STAI score.
- Neuroticism, EDS, extraversion, agreeableness, STAI and conscientiousness present in decreasing order the highest dispersion between the two groups in RF.
- The highest separation between the two groups in LDA is given by EPDS

Model comparisson

Table 2. Overall results and comparative statistic of the methods

Method	Misclassification Matrices		Overall Accuracy	Sens.	Spec.	AUC
		Actual				
LR	Predicted	586 36 151 70	0.78	0.8	0.66	0.74
CT	Predicted	724 85 13 21	0.88	0.97	0.25	0.59
RF	Predicted	723 84 14 22	0.88	0.98	0.21	0.8
LDA	Predicted	715 76 22 30	0.88	0.97	0.28	0.83

Figure 4. Resulting ROC curves



Conclusions

- The overall accuracy of the models suggests that the non-parametric methods of recursive partitioning and the dimension reduction method of LDA perform comparatively better than LR.
- LR, however, maintains a higher specificity value.

Limitations

- As this is a work in progress, all models have not been optimized.
- Bias and variance have not been accounted for.
- Improve the imputation method of the missing classes

Reference: Meijer, J. L., Beijers, C., Van Pampus, M. G., Verbeek, T., Stolk, R. P., Milgrom, J., Bockting, C. L.H., Burger, H. (2014). Predictive accuracy of Edinburgh postnatal depression scale assessment during pregnancy for the risk of developing postpartum depressive symptoms: A prospective cohort study. *BJOG: An International Journal of Obstetrics and Gynaecology*, *121*(13), 1604–1610. https://doi.org/10.1111/1471-0528.12759

Corresponding author: s.ioannou@umcg.nl