

The role of machine learning in identification of early gestational diabetes mellitus prediction models

Authors: Sekulić M¹, Relić D^{1,2}(mentor)

¹ School of medicine, University of Zagreb, Zagreb, Croatia

² School of Public Health "Andrija Štampar", Zagreb, Croatia

DOI: <https://doi.org/10.26800/LV-142-sup5-61>

Introduction:

Gestational diabetes mellitus(GDM) is a condition in which carbohydrate intolerance develops during pregnancy. The estimated prevalence of GDM ranges from less than 1% to 28% and is commonly diagnosed between 24 and 28 weeks of gestation. Untreated GDM represents a severe threat to the affected women and their offspring. Machine learning (ML) is a computer science discipline focused on algorithms that improve automatically through experience to make predictions or decisions without being explicitly programmed to do so.

Aim:

Review of machine learning generated models for early screening of GDM and identification of possible novel research points for future works.

Materials & Methods:

The search using "gestational diabetes mellitus" or "gestational diabetes" as a part of the title and "machine learning" as a topic was conducted on the 7th of October through Web of science core collection bases and yielded 15 results from which 8 were included in this paper. Only full-text articles that examined early diagnosis of gestational diabetes were included.

Results:

All studies established models for early prediction of GDM, studies included were prospective cohort, retrospective cohort, and case-control studies. One study focused only on obese women. Two studies examine microRNA and visceral fat mass as novel biomarkers for GDM prediction. Four studies showcase recommended division of data into training(2/3) and validation(1/3) datasets.

Conclusion:

As the prevalence of GDM increases, due to the limited impact of therapeutic measures in the second and third trimester, an even greater need for novel predictors in earlier pregnancy arises. ML offers wide possibilities for data interpretation in a non-linear manner. Drawbacks of proposed models originating from current studies are lack of cross-validation and smaller data sample size in some studies where the full extent of ML cannot be observed. A further meta-analysis should be conducted to estimate the significance of the studies.

Keywords: gestational diabetes mellitus, machine learning, prediction model