Mahidol University
Faculty of Medicine Ramathibodi Hospital

Mixed effect machine learning:

A framework for predicting longitudinal change in hemoglobin A1c

Commentary by

Pongsathorn Piebpien

Doctor of Philosophy Program in Data Science for Health Care

Department of Clinical Epidemiology and Biostatistics

Mahidol University Faculty of Medicine Ramathibodi Hospital

Question from the topic

Are there any study about machine learning to predict outcome in longitudinal study?

Does mixed effect machine learning model is the best choice for predicting longitudinal outcome?

Artificial Intelligence In Medicine 98 (2019) 109–134

Contents lists available at ScienceDirect

Artificial Intelligence In Medicine

journal homepage: www.elsevier.com/locate/artmed

Data-driven modeling and prediction of blood glucose dynamics: Machine learning applications in type 1 diabetes

Ashenafi Zebene Woldaregay[a,*], Eirik Årsand[b], Ståle Walderhaug[b,c], David Albers[d], Lena Mamykina[d], Taxiarchis Botsis[e], Gunnar Hartvigsen[a]

[a] Department of Computer Science, University of Tromsø-The Arctic University of Norway, Tromsø, Norway
[b] Norwegian Centre for E-health Research, University Hospital of North Norway, Tromsø, Norway
[c] SINTEF Digital, Software Engineering, Safety and Security, Tromsø, Norway
[d] Department of Biomedical Informatics, Columbia University, N.Y., USA
[e] The Sidney Kimmel Comprehensive Cancer Center, Johns Hopkins University School of Medicine Baltimore, MD, United States

About this article

- Introduction of machine learning and its application for intelligent and improved systems for diabetes management
- Modeling options and strategies of machine learning and a hybrid system focusing on the prediction of BG dynamics in type 1 diabetes
- Literature review (Final: 55 articles)
  - review is mainly on BG prediction techniques
  - classes of machine learning algorithms, artificial neural networks, support vector machines, Bayesian neural networks, decision trees, and others
  - reinforcement learning is not under the scope of this review

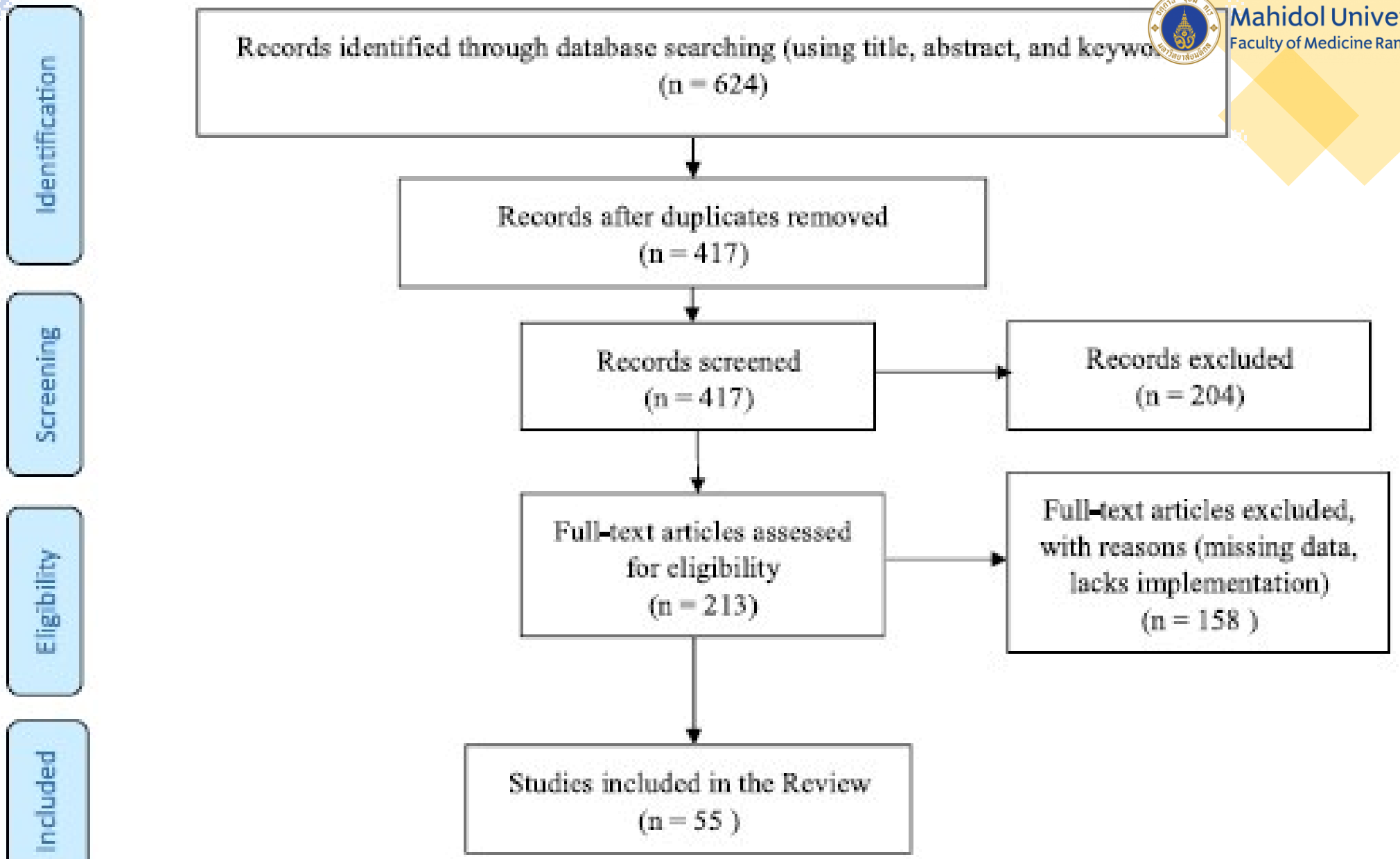


Fig. 5. Diagram of the review process.

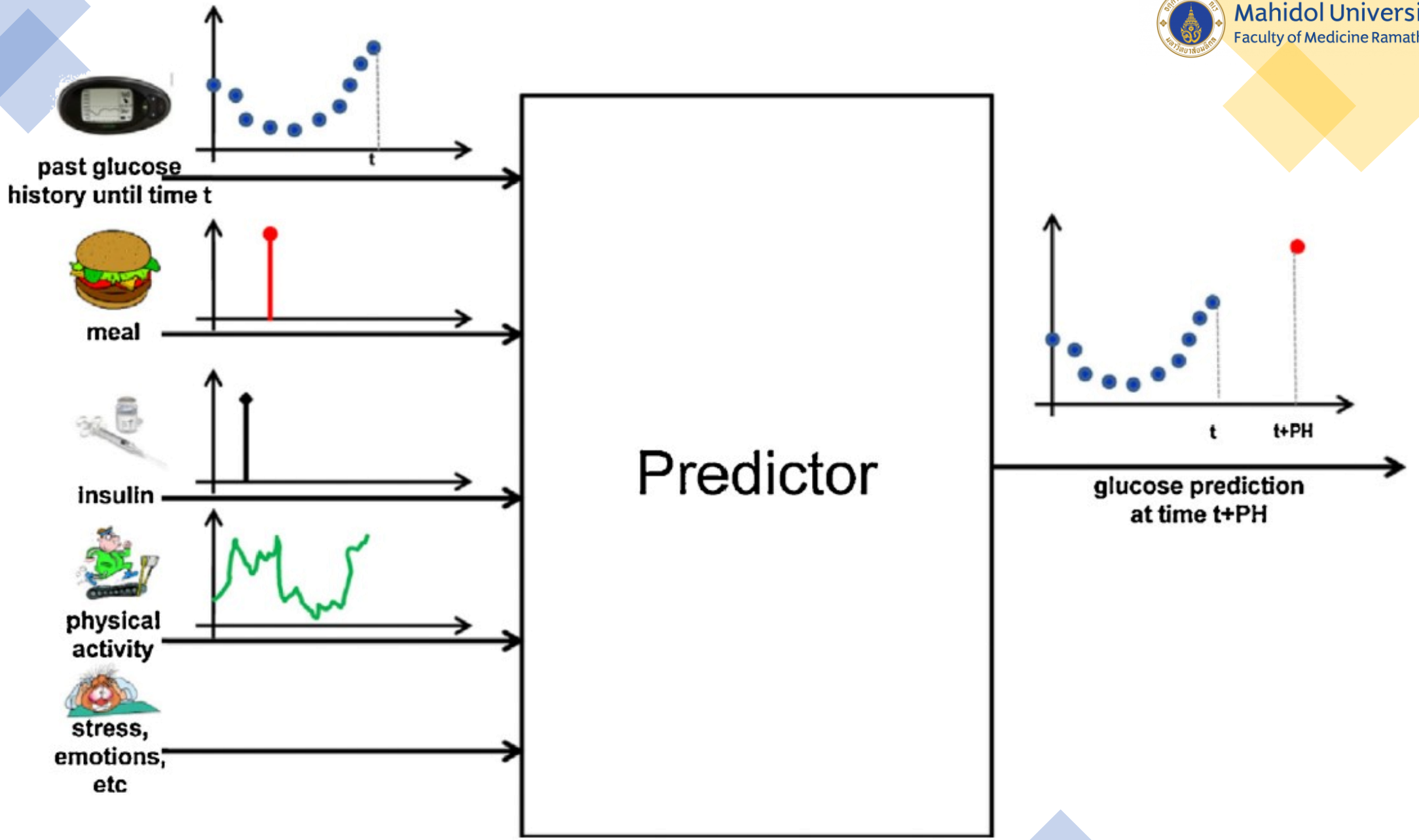


Fig. 2. An ideal blood glucose predictor (Reprinted from [15]).

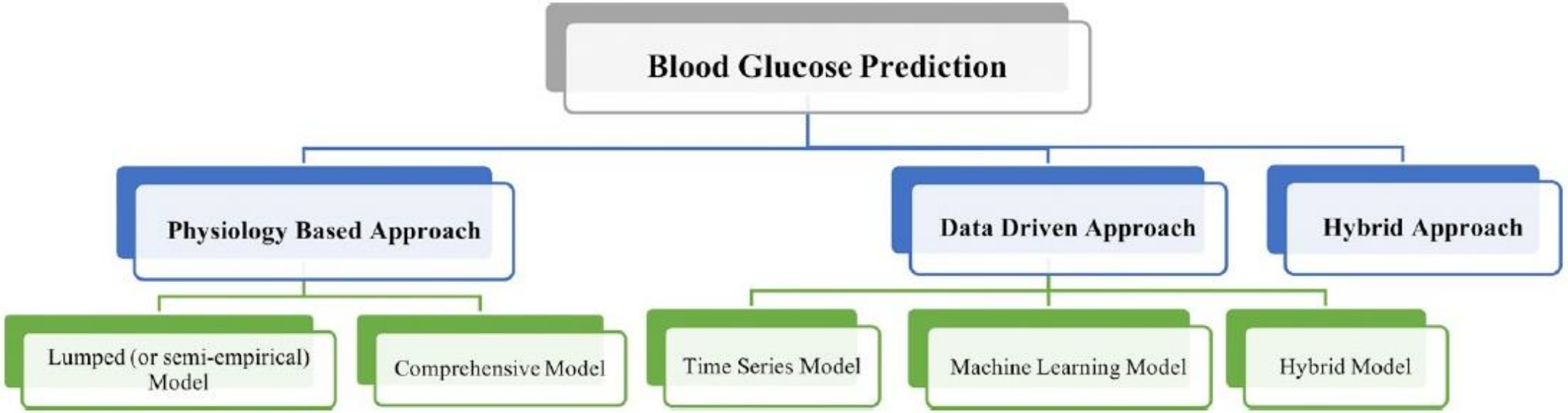


Fig. 3. Taxonomy of blood glucose prediction approaches.

Summary of existing efforts (Machine learning techniques)

- *Artificial neural network (ANN)*
- *Support vector machines (SVM), kernel function (KF), and gaussian process regression*
- *Genetic programming and genetic algorithms*
- *Random Forest (RF)*
- *Hybrid approach*
- *Ensemble approach – merging different predictors for performance improvement*

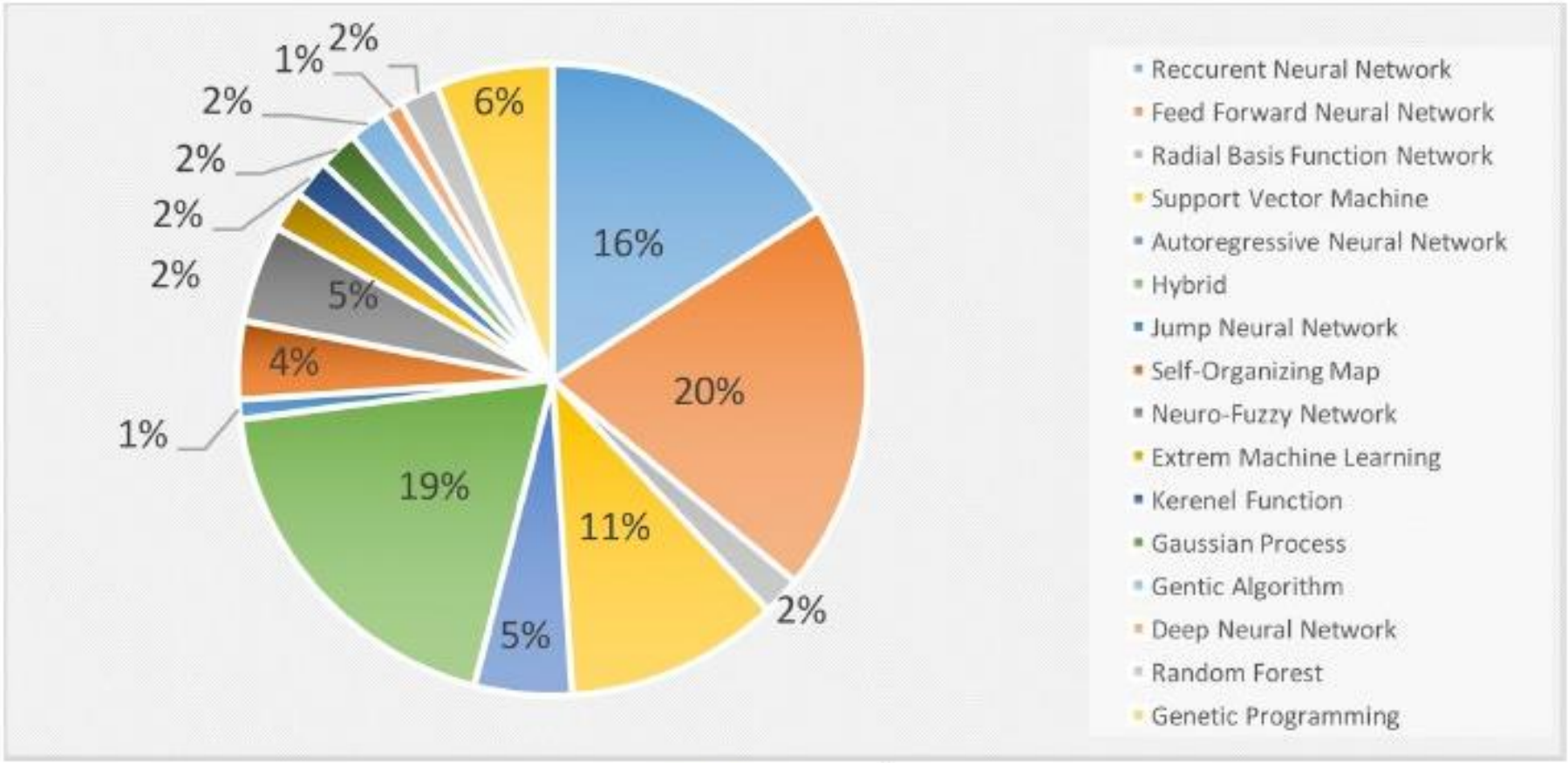


Fig. 7. Classes of machine learning techniques used in the modeling of blood glucose prediction. (For interpretation of the references to colour in this figure, the reader is referred to the web version of this article).

Classes of machine learning

Model Performance evaluations

Performance metrics are necessary steps that should be carefully chosen based on the developed model under consideration

Various performance metrics are used to assess the predictive power of the developed model

- mathematical evaluation criteria (empirical accuracy)
- clinical evaluation criteria (clinical accuracy)

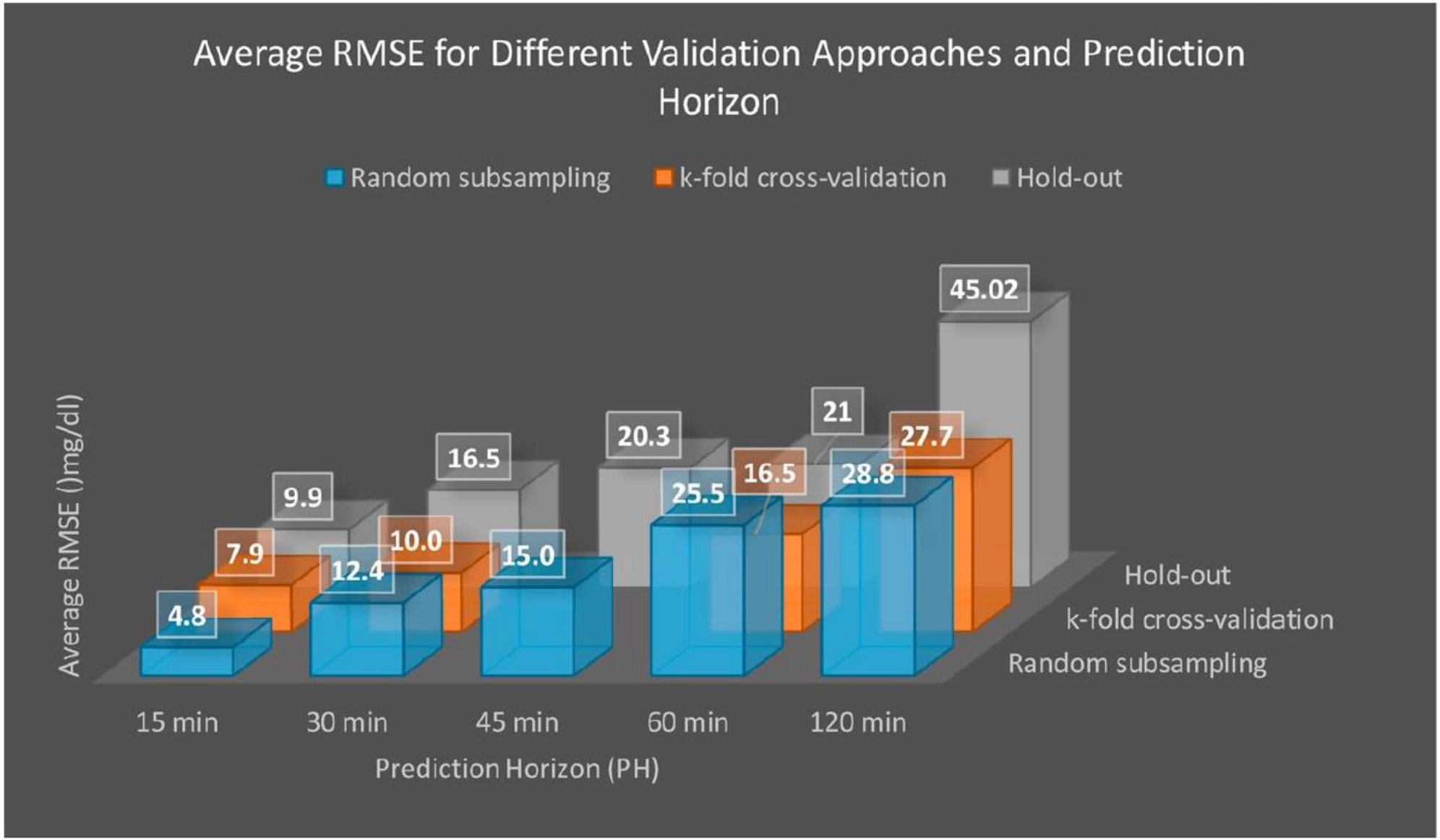
Prediction horizon (PH) is the lead time in which the model is able to forecast BG levels in the future.

A comparative assessment of prediction performances

- The predictive performance of a BG predictor is affected by many technical factors and design choices, apart from the factors that affect BG dynamics, such as the type of machine learning, data size, prediction horizon (PH) and validation approaches considered, and others
- The most popular performance metrics, RMSE, to compare the reported performance levels by different researchers

Validation strategies

- complementary issues
  - Overfitting
  - Under training
- Enough data size in the three datasets (training, testing, and validation sets)
  - Proportion of splitting
  - The most popular: k-fold cross-validation
  - K-fold cross-validation strategy involves splitting the datasets into randomly partitioned k equal subsets and using one set as a validation sets and the rest for training, repeating the same process for all the different subsets
  - Hold-out involves portioning the datasets into non-overlapping subsets, where the first subset is entirely used for training and the rest for testing


Average RMSE for Different Validation Approaches and Prediction Horizon
Random subsampling
k-fold cross-validation
Hold-out
Average RMSE ()mg/dl)
4.8
7.9
9.9
12.4
10.0
16.5
15.0
20.3
25.5
16.5
21
28.8
27.7
45.02
Hold-out
k-fold cross-validation
Random subsampling
15 min
30 min
45 min
60 min
120 min
Prediction Horizon (PH)

Findings from this articles

machine learning has received wider attentions for modeling and the prediction of BG dynamics

this review is to assess and analyze the machine learning applications in BG prediction

it remains difficult to achidue to the complexity of BG dynamics, eve an accurate prediction

The reported BG prediction algorithms have explored various classes of machine learning, input parameters, and training algorithms

most of the algorithms rely on in silico evaluations

clinical significance of model still in question?

Clinical significance of model in question?

Most researchers have considered separate age groups, which are typically related with the dynamics and active lifestyles adopted by each group

Most of the reported algorithms perform better in either of these BG regions (hypo/eu/hyperglycemia events)

Most of model lack of a well-defined variables can severely affect the predictive performance.

Answer the question from the topic

Are there any study about machine learning to predict outcome in longitudinal study?

Does mixed effect machine learning model is the best choice for predicting longitudinal outcome?