

# Federated Learning using IoT Devices

Rishabh Anand  
CSE-IoT  
Chandigarh University  
Mohali, Punjab  
19BCS4525@cuchd.in

Udita Mitra  
CSE-IoT  
Chandigarh University  
Mohali, Punjab  
19BCS4662@cucd.in

Khushwant Singh Rathore  
CSE-IoT  
Chandigarh University  
Mohali, Punjab  
19BCS4644@cuchd.in

Vandana Chauhan  
CSE-IoT  
Chandigarh University  
Mohali, Punjab  
19BCS45322@cuchd.in

Abhishek Gupta  
CSE-IoT  
Chandigarh University  
Mohali, Punjab  
19BCS4579@cuchd.in

**Abstract** - The advances of sensing and computing technologies pave the way to develop novel applications and services for wearable devices. For example, wearable devices measure heart rate, which accurately reflects the intensity of physical exercise. Therefore, heart rate prediction from wearable devices benefits users with optimization of the training process. Conventionally, Cloud collects user data from wearable devices and conducts inference. However, this paradigm introduces significant privacy concerns. Federated learning is an emerging paradigm that enhances user privacy by remaining the majority of personal data on users' devices. In this paper, we propose a statistically sound, Bayesian inference federated learning for heart rate prediction with autoregression with exogenous variable (ARX) model. The proposed privacy-preserving method achieves accurate and robust heart rate prediction. To validate our method, we conduct extensive experiments with real-world outdoor running exercise data collected from wearable devices.

**Index Terms**—Federated learning, Bayesian inference, Wearable computing, Heart rate prediction.

## I. INTRODUCTION

Cardiovascular diseases are the top cause of deaths globally. According to the WHO report, 17.9 million people die from CVD each year, an estimated 31% of all deaths worldwide [1]. Many factors can trigger these diseases. Preventing CVD is becoming an important task. It is already known that exercising has a proven therapeutic effect on the cardiovascular system. Therefore, predicting and controlling heart rates during the exercise is becoming more and more important to avoid overstrain and prevent sudden heart-rate break.

Wearable devices enable smart human-computer interactions. The wearable fitness/sport industry is expected to grow exponentially in the near future. Users of wearable devices expect a service that can guide their smart exercise coaching, rather than only tracking their activities. Heart rate based training is a good technique to improve the effectiveness of training and prevent over-exercising. Designing an optimal exercise training plan to avoid overstrain is crucial. Ignoring the limits of the physical activities will not only nullify the effect of the exercise but also cause harmful effects on the cardiovascular system. The very first step of designing the optimal exercise training plan is predicting heart rate from the exercise, which will then be used for the training control. The designed recommendation and control systems (an after effect of the whole project) is displayed on the site.

Subjects can use the control system in those smart devices to guide their exercise in order to reach the desired heart rate response and avoid overtraining, which will benefit the users' health. Most of the existing research work related to heart rate prediction focuses on indoor exercises. But for outdoor physical exercise, it is not possible to automatically regulate the workload intensity due to the dependence on environmental conditions. Therefore, it's typical for an outdoor exerciser to continuously check heart rate and increase or decrease the speed accordingly for regulating his or her heart rate.

Machine learning has demonstrated its promising performance in providing the users with recommendations regarding physical activity and physiological response [2, 3]. Machine learning algorithms typically learn from centralized data in order to train a powerful model. But, pooling data from many users to the Cloud introduces significant privacy concerns; for example, leaking sensitive health information of users. Recently, EU General Data Protection Regulation (GDPR) [4] states the need for trust to be built into personal data services and allows users to control their own data, including data their devices generate.

Our contributions are threefold:

- We propose a federated learning method based on the Sequential Bayesian method (FD Seq Bayes), for heart rate prediction without pooling data to the Cloud for privacy preservation. The model is proposed to provide a statistically sound way of integrating local models
- We have conducted extensive evaluation on real-world data from wearable devices that we built ourselves using

ESP32 and MAX30100/02. Compared to various state-of-the-art baseline models, our proposed methods have demonstrated their strength in achieving higher prediction accuracy on unseen, new users with lower computation cost.

- Our proposed Bayesian federated learning methods can be easily extended to address other ARX regression problems taking consideration of user privacy preservation and achieving good performance.

## II. PROPOSED APPROACH

This section presents the problem statement on federated learning for heart rate prediction and introduces Bayesian-based technique sequential models.

### A. Problem Definition

The objective of heart rate prediction is to predict the heart rate  $y_t$  given the historic readings of the previous heart rates and other useful inputs like speed:

$$y_t = f(y_{1:t-1}, x_{1:t}) + e_t$$

where  $e_t$  is the error term and  $f$  can be any parametric function, say a linear function or neural network. The objective of federated learning is to learn such a parametric model  $f$  in the server without sending each user's raw data. In particular, given data from  $n$  different users stored at each distributed node, and denote the reading from user  $i$  as  $D_i$ , the learning outcome is a trained global model in the server with datasets  $D_i^n = 1$  and the global learning should only involve model parameters rather than raw user data. For later prediction, the trained model at the server can then be directly used for predictions of future users with personalization if possible.

### B. Autoregression with Exogenous Variable Model

As the heart rate data is a time series with serial correlations, a suitable model for such data sets is ARX. An ARX model with  $p$  autoregression components and  $q + 1$  lagged inputs can be formally written as:

$$y_t = \theta_0 + \sum_{i=1}^p \theta_i y_{t-i} + \sum_{j=0}^q \omega_j z_{t-j} + e_t$$

where  $e_t \sim N(0, \sigma^2)$  is white noise with variance  $\sigma^2$ ,  $y_t, z_t$  are heart rate and speed measurements at time  $t$ . By defining  $\beta, x$  as the vectors concatenating the model parameters and covariates, the model can be succinctly written as:

$$y_t = x^T \beta + e_t$$

where  $\beta^T = [\theta_0, \theta_1, \dots, \theta_p, \omega_0, \dots, \omega_q]$  and  $x^T = [1, y_t - 1, \dots, y_{t-p}, x_t, \dots, x_{t-q}]$ .

### C. Federated Learning with Sequential Bayesian Inference

Bayesian inference provides a natural solution to the federated learning problem, where the inference is on the posterior distribution of model parameters. By making conditional independent assumption of the data at different nodes given the model parameter, the posterior distribution of the model parameter can be learnt in a sequential manner.

Denoting  $D_i = X_i, y_i$  as the dataset at node  $i$  where  $y_i = [y_{i,1}, \dots, y_{i,n_i}]^T$  and  $X_i = [x_{i,1}, \dots, x_{i,n_i}]^T$ , and  $n_i$  is the number of time instances for user  $i$ ; then the posterior distribution is

$$\begin{aligned} & p(\beta, \sigma^2 | D_1, D_2, \dots, D_n) \\ & \propto p(\beta, \sigma^2) p(D_1, D_2, \dots, D_n | \beta, \sigma^2) \\ & = p(\beta, \sigma^2) \prod_{i=1}^n p(D_i | \beta, \sigma^2) \\ & \propto p(\beta, \sigma^2 | D_1, D_2, \dots, D_{n-1}) p(D_n | \beta, \sigma^2) \\ & i = 1 \end{aligned}$$

## III. RESULTS

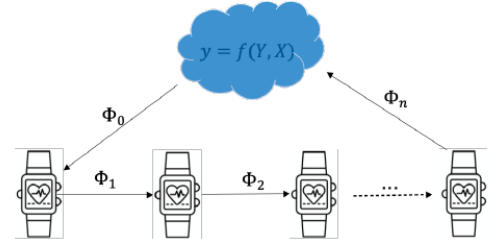
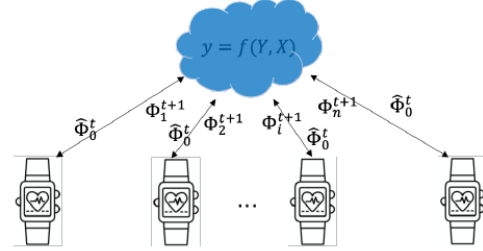
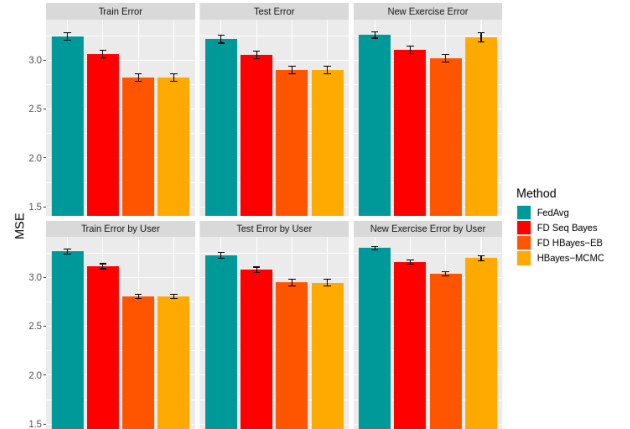
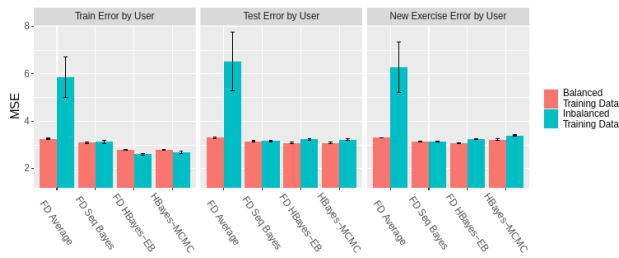


Fig. 1: Federated learning with sequential Bayesian inference



$$\begin{aligned} \text{error}_{\text{time}} &= \frac{\sum_{i=1}^n \sum_{t=1}^{n_i} (y_{i,t} - \hat{y}_{i,t})^2}{\sum_{i=1}^n n_i}, \\ \text{error}_{\text{user}} &= \frac{\sum_{i=1}^n (\sum_{t=1}^{n_i} (y_{i,t} - \hat{y}_{i,t})^2 / n_i)}{n}, \end{aligned}$$





#### IV. CONCLUSION

When users perform physical exercises, one important goal is to optimize the training process. Heart rate has been used as a most important indicator for monitoring the training strain. Therefore, predicting heart rate during physical exercise is crucial for tracking physiological responses and improving the effect of the exercise. The majority of the existing research focuses on pooling together a large amount of users' data for building a robust model, which often has incurred much privacy concern. To tackle this issue, we leverage a statistically sound model – Bayesian inference and propose a Bayesian-based federated learning method i.e. FD Seq Bayes. They enable collaborative model training under the orchestration of a central server, while not accessing to any user's local data. Through extensive evaluation on real-world dataset, we have demonstrated the advantages of our methods in accurate prediction and low computation cost. In the future, we will extend our evaluation to other ARX regression problems to assess the generalization of our methods.

#### V. REFERENCES

- 1) Hilmkil, A., Ivarsson, O., Johansson, M., Kuylensstierna, D., Erp, T. V.: Towards Machine Learning on data from Professional Cyclists. In: 12th World Congress on Performance Analysis of Sports, Opatija, Croatia (2018).
- 2) Cheng, T. M., Savkin, A. V., Celler, B. G., Su, S. W., Wang, L.: Nonlinear modeling and control of human heart rate response during exercise with various work load intensities. *IEEE Trans. Biomed. Eng.* 55(11), 2499–2508 (2008).
- 3) Mohammad, S., Guerra, T. M., Grobois, J. M., Hecquet, B.: Heart rate control during cycling exercise using Takagi-Sugeno models. In: 8th IFAC World Congress, pp. 12783–12788, Milano, Italy (2011).
- 4) Ni, J. M., Muhlstein, L., McAuley, J.: Modeling Heart Rate and Activity Data for Personalized Fitness. In: WWW'19, May 13–17, 2019, San Francisco, CA, USA.
- 5) Kairouz, P., et al.: Advances and Open Problems in Federated Learning. *arXiv:1912.04977* (2019).
- 6) Konečný J., McMahan H. B., Ramage D., Richtárik P.: Federated Optimization: Distributed Machine Learning for On-Device Intelligence. *arXiv:1610.02527* (2016).
- 7) Su, S. W., Wang, L., Celler, B. G., Savkin, A. V., Guo, Y.: Identification and control for heart rate regulation

during treadmill exercise. *IEEE Trans. Biomed. Eng.* 54(7), 1238–1246 (2007b).

- 8) Ludwig, M., Hoffmann, K., Endler, S., Asteroth, A., Wiemeyer, J.: Measurement, Prediction, and Control of Individual Heart Rate Responses to Exercise-Basics and Options for Wearable Devices. *Front. Physiol.* (2018). <https://doi.org/10.3389/fphys.2018.00778>