# PREgDICT : Early prediction of gestational weight gain for pregnancy care

Chetanya Puri[1,2+], Gerben Kooijman[2], Felipe Masculo[2], Shannon Van Sambeek[2], Sebastiaan Den Boer[2], Stijn Luca[3*], and Bart Vanrumste[1*]

*Abstract*— **Excessive or inadequate Gestational Weight Gain (GWG) is considered to not only put the mothers, but also the infants at increased risks with a number of adverse outcomes. In this paper, we use self-reported weight measurements from the early days of pregnancy to predict and classify the end-of-pregnancy weight gain into an underweight, normal or obese category in accordance with the Institute of Medicine recommended guidelines. Self-reported weight measurements suffer from issues such as lack of enough data and non-uniformity. We propose and compare two novel parametric and non-parametric approaches that utilise self-training data along with population data to tackle limited data availability. We, dynamically find the subset of closest time series from the population weight-gain data to a given subject. Then, a non-parametric Gaussian Process (GP) regression model, learnt on the selected subset is used to forecast the self-reported weight measurements of given subject. Our novel approach produces mean absolute error (MAE) of 2.572 kgs in forecasting end-of-pregnancy weight gain and achieves weight-category-classification accuracy of 63.75% mid-way through the pregnancy, whereas a state-of-the-art approach is only 53.75% accurate and produces high MAE of 16.22 kgs. Our method ensures reliable prediction of the end-of-pregnancy weight gain using few data points and can assist in early intervention that can prevent gaining or losing excessive weight during pregnancy.**

## I. INTRODUCTION

Global trends suggest [1] that around 47% of the pregnant women end-up being overweight or obese at the end of gestational period and around 23% tend to gain too little weight during their pregnancy. Gaining too much or too little weight during the pregnancy can cause several short-term and long-term health related problems not only to the mother but also the infant. Excessive Gestational weight Gain (GWG) can lead to postpartum weight retention and subsequent maternal obesity [2] that increases the risk of gestational diabetes [3]. It can also result in large-for-gestational-age infant and/or caesarean delivery or other labor and delivery complications [1]. Foreseeable risks for infants such as childhood obesity are also reported [4]. On the contrary, gaining too little weight during pregnancy is also not considered healthy. This increases the risk of preterm birth or small-for-gestational-age infants [1]. Institute of Medicine (IOM) recommends a set of guidelines [5] on how much weight women in different body mass index (BMI) categories should gain during their pregnancy to deliver at a normal weight (Table I).

[1]Campus Groep T, e-Media Lab, KU Leuven, Belgium
[2]Philips Research, Eindhoven, The Netherlands
[3]Department of Data Analysis and Mathematical Modelling, Ghent University, Gent, Belgium
[+]email : chetanya.puri@kuleuven.be
[*]These authors contributed equally to this work.

Most of the women do not follow the guidelines or realize when it's too late into the pregnancy. Our proposed solution can accurately predict the trend in weight gain using the weights during the initial days of the pregnancy. This can help prenatal care providers in risk assessment during a pregnancy and provide better adaptive coaching to the mothers. Moreover, mothers can track the *rate* of weight gain and use our proposed model to monitor weight gain, thus reducing GWG related risks at the end of their pregnancy.

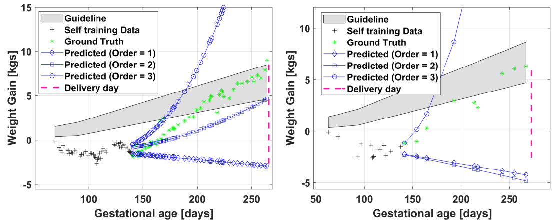
We use *self-reported* real life weight measurements data from 80 expecting mothers collected during their pregnancy. We formulate this as an absolute weight prediction problem with the end goal of predicting the weight at the end of the pregnancy and classifying if the weight is within the guidelines or not. We have restricted our analysis to the mothers with singleton pregnancy for this study. Data from mothers expecting more than one child is very rare to obtain. Also, the guidelines [5] for gestational weight gain consider singleton mothers.

TABLE I: 2009 IOM guidelines for weight gain and rate of weight gain during pregnancy

| Pre-pregnancy Body Mass Index (BMI) category | Mothers of singletons | |
|---|---|---|
| | Total weight gain (in kgs) | Rate of weight gain in the second and third trimesters (kg/wk) |
| Underweight (<18.5 kg/m2) | 12.7-18.14 | 0.45-0.59 |
| Normal-weight (18.5-24.9 kg/m2) | 11.34-15.88 | 0.36-0.45 |
| Overweight (25.0-29.9 kg/m2) | 6.8-11.34 | 0.23-0.32 |
| Obese ($\geq$ 30.0 kg/m2) | 4.99-9.07 | 0.18 -0.27 |

Weight gain during pregnancy varies from person to person based on their pre-pregnancy weight and BMI ranges. Women with underweight pre-pregnancy BMI are expected to gain weight at a higher rate than women that were overweight before pregnancy. If enough reliable weight measurements collected uniformly over time are available for training, one can estimate the end-of-pregnancy weight gain by fitting a first to third order polynomial. However, there are two major challenges i) weight measurement data suffers from noise, is incomplete, sparse and non-uniformly sampled due to the self-reported nature, ii) limited availability of training data from the initial few days of the pregnancy for early prediction. Fig. 1 shows weight gain measurements and early prediction using polynomial fitting from two subjects.

In this paper, we experiment and compare *parametric* bayesian regression and *non-parametric* gaussian process regression to model the time series data. We find a subset of population data that is close to our test subjects' limited




(a) Good self-training data (Less sparse, pseudo-uniformly sampled), $i = 9$

(b) Bad self-training data (Highly sparse, non-uniformly sampled), $i = 35$

Fig. 1: *Self-reported* weight gain data are sparse, non-uniformly sampled. Forecasting for $i^{th}$ subject on delivery day by fitting $1^{st}$ to $3^{rd}$ order polynomial is highly inaccurate for early prediction (with limited self-training data)

self-training data. We then combine models learnt from this population subset along with personal data for early prediction of gestational weight gain. Autoregressive integrated moving average (ARIMA) models [6] are considered as a state-of-the-art approach in time series forecasting that learn structures from the time series data, given sufficient historical personal data. We show that our approach out-performs state-of-the-art in early prediction by tuning the personal model using the general population model learnt a-priori.

## II. METHODOLOGY

We are given a population gestational weight gain time series data from $N$ subjects as $\mathcal{X} = \{(\mathbf{x}^1, \mathbf{y}^1), ...(\mathbf{x}^N, \mathbf{y}^N)\}$, where $\mathbf{x}^i = \{t_1^i, t_2^i, t_3^i, ..., t_m^i\}$ represents the input gestational days upto delivery day $t_m^i$ and $\mathbf{y}^i = \{y_1^i, y_2^i, y_3^i, ..., y_m^i)\}$ represents the output weight gain for $i^{th}$ subject, where $y_k^i = y(t_k^i)$. It is important to note here that $t_1^i$ *does not necessarily equal* $t_1^j, i, j \in \{1, 2, ...N\}$. This is because the data is *self-reported* and each time of measurement varies from user to user according to their personal preferences and adherence to data collection.

Additionally, we are given individual weight measurements from a subject's initial $t_d$ days of pregnancy data, $\mathcal{D} = \{(t_1^+, y_1^+), (t_2^+, y_2^+)..., (t_d^+, y_d^+)\}$. We call this the self-training data. Weight gain data from population data over entire gestational period is called population-training data.

The objective is to try to learn function(s) $f$ from given population and individual data, such that,

$$y^+ = f(t^+) + \epsilon \tag{1}$$

where $\epsilon \sim \mathcal{N}(0, \sigma^2)$ is independent and identically distributed (i.i.d) gaussian.

Our parametric approach learns parameters' information a-priori from the population data and the non-parametric approach learns closest subset from the population data. We then use this knowledge to build personalised models and learn $f$. The individual weight gain in future at delivery time $t_m^+$ is forecasted using the learned model $f$ and $y_m^+ = y(t_m^+) = f(t_m^+)$. Next, we discuss these approaches in detail.

### A. Parametric

We can fit a $p^{th}$-order polynomial with $f = w_0 + w_1 t + w_2 t^2 + ... + w_p t^p$ in eq. (1) and estimate the coefficients $\mathbf{w} = [w_0, w_1, ...w_p]^T$ by maximizing the likelihood over individual's self-training data $\mathcal{D}$, $\mathcal{L}(\mathbf{w}) = P(\mathcal{D}|\mathbf{w})$,

$$\hat{\mathbf{w}}_{MLE} = \operatorname*{argmax}_{\mathbf{w}} P(\mathcal{D}|\mathbf{w}) = \prod_{i=1}^{d} p(y_i^+|t_i^+; \mathbf{w}) \tag{2}$$

Eq. (2) refers to the model learnt from the individual's sparse limited observations data upto given $t_d$ days. Next, we exploit the population-training data and find the maximum likelihood estimates of $\hat{\mathbf{w}}$ for all the time series in the population-training data. We, then use the distribution of the estimates of these coefficients, $p(\mathbf{w})$ acquired from the $N$ subjects in the population data as an *a-priori* estimate. The likelihood learnt from the self-training data and the *a-priori* distribution learnt from the population data are then combined using bayes theorem to calculate the maximum-a-posteriori estimate of the coefficients $p(\mathbf{w}|\mathcal{D})$.

$$\hat{\mathbf{w}}_{MAP} = \operatorname*{argmax}_{\mathbf{w}} p(\mathbf{w}|\mathcal{D}) = \frac{P(\mathcal{D}|\mathbf{w})p(\mathbf{w})}{P(\mathcal{D})} \tag{3}$$

The forecast at time $t_m^+$ is given by $\hat{\mathbf{w}}_{MAP}[t_m^+ \, {t_m^+}^2 .... {t_m^+}^p]^T$. We call this approach parametric because the choice of order of the polynomial $p$ depends on the application of interest, which should be pre-defined.

### B. Non-Parametric

In this subsection, we discuss Gaussian Processes (GP) as a non-parametric approach that is robust to sparse and limited availability of data, especially in medical time series data such as vital signs monitoring [7]. GP is defined as a set of random variables, such that any finite number of them have a joint Gaussian distribution [8]. Here, we define '$f$' from eq. (1) as a GP specified by mean and covariance functions $m(t)$ and $k(t, t')$ respectively, $f(t) \sim \mathcal{GP}(m(t), k(t, t'))$. The covariance function models the structure of the time series by exploiting the relationship between two independent observations based on our assumptions of data, given that the covariance remains positive semi-definite [8]. Assuming that the data is noisy with i.i.d gaussian noise, having noise covariance $\sigma_n^2$, we chose a gaussian covariance function,

$$k(t, t') = \sigma_f^2 \, exp \left[ \frac{-(t - t')^2}{2l^2} \right] \tag{4}$$

We chose a Gaussian covariance function as prior because it is almost unity when two observations are very close in time, and decreases as they become far apart in time. This is in line with the fact that gestational weight gain does not changes abruptly and is correlated in time.

Let's assume, we find a subset $\hat{\mathcal{X}}$ from the given population dataset $\mathcal{X}$, $\hat{\mathcal{X}} = \{(\mathbf{x}^1, \mathbf{y}^1), ..., (\mathbf{x}^M, \mathbf{y}^M)\} = \{(t_1^1, y_1^1), ..., (t_m^1, y_m^1), ..., (t_1^M, y_1^M), ..., (t_m^M, y_m^M)\}$, $M << N$ using some *heuristic* such that $\hat{\mathcal{X}}$ is closer to the individual test subject's data. We discuss the computation and advantages of this *heuristic* later in the section II-B.1.

Given $\hat{\mathbf{y}} = [y_1^1, ..., y_m^1, ...., y_1^M, .., y_m^M]^T$ and $\mathbf{K}$ as a matrix of entries $K_{ab} = k(t_a, t_b), \forall t_a, t_b \in \hat{\mathcal{X}}$ using eqn. (4), we follow optimisation procedure from [8] to estimate the hyperparameters $\{\sigma_f,\ l,\ \sigma_n\}$ by maximising the marginal likelihood $p(\hat{\mathbf{y}}|\hat{\mathcal{X}}; \{\sigma_f,\ l,\ \sigma_n\})$. The prediction at time $t_m^+$ is given as a gaussian distribution whose mean, $\mu$ and variance, $\sigma^2$ are given by $\mu(t_m^+) = \mathbf{k}_+{}^T(\mathbf{K} + \sigma_n^2\mathbf{I})^{-1}\hat{\mathbf{y}}$, $\sigma(t_m^+) = k(t_m^+, t_m^+) - \mathbf{k}_+{}^T(\mathbf{K} + \sigma_n^2\mathbf{I})^{-1}\mathbf{k}_+$, where $\mathbf{k}_+ = \mathbf{k}(t_m^+)$, $\mathbf{k}(t_m^+) = [k(t_m^+, t_1^1), ..., k(t_m^+, t_m^M)]^T$

*1) Heuristics to find $\hat{\mathcal{X}}$ from $\mathcal{X}$:* We find the subset $\hat{\mathcal{X}}$ from $\mathcal{X}$ such that elements in $\hat{\mathcal{X}}$ are in some sense closest to the individual self-training data, $\mathcal{D}$. This gives an advantage in terms of prediction capability and reduces the computational complexity. It improves the prediction capability by reducing the variability at time $t$ in training data occurring due to inter-personal variations among subjects from complete population and selecting only those subjects, who exhibit similar weight gain trend. The computational complexity of Gaussian processes is $\mathcal{O}(n^3)$, where $n$ is the number of training points in a dataset. Clearly, $n(\hat{\mathcal{X}}) << n(\mathcal{X})$ because $M << N$. Although, our heuristics method give a relative computational advantage over vanilla gaussian process methods, still the computational complexity is expected to be too high if $N$ is of the order of $10^6$, which is a limitation of this non-parametric approach. Sparse GPs detailed in [8] are model approximation methods that can further reduce the computational complexity, if scaling to such high number of subjects.

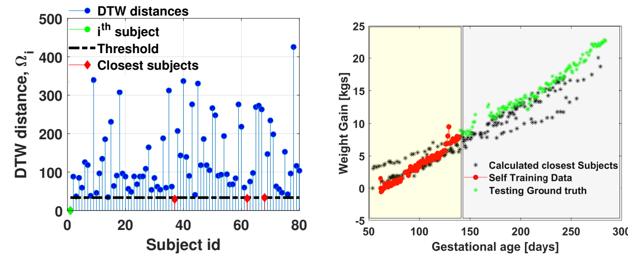
In order to find $\hat{\mathcal{X}}$ from $\mathcal{X}$, first, we calculate the distances between $i^{th}$ test subject's self-training data $\mathcal{D} = \{(t_1^+, y_1^+), (t_2^+, y_2^+)..., (t_d^+, y_d^+)\}$ and given population subjects' data (until allowed training day '$t_d^+$') as $\mathbf{\Omega_i} = [\omega_{i1}, \omega_{i2}, ....\omega_{iN}]^T$, where $\omega_{ij} = dist([y_1^+, ...y_d^+], [y_1^j, ...y_d^j])$. We have chosen Dynamic time warping (DTW) [9] as distance measure $dist$ in this work due to it's robustness over other distance measures in indexing time series [10] and it's capability to handle non-uniformly sampled data. Next, we actively find the closest subjects by calculating cut-off point for the $i^{th}$ subject as follows; i) arrange the distance vector $\mathbf{\Omega_i}$ in ascending order to calculate $\hat{\mathbf{\Omega}}_\mathbf{i} = [\hat{\omega_{i1}}, \hat{\omega_{i2}}, ....\hat{\omega_{iN}}]^T$, such that $\hat{\omega_{ik}} \leq \hat{\omega_{ik+1}} \forall k = 1, 2, ..N$, ii) find *turning* points at location '$k$', such that, $(\hat{\omega}_{ik-1} - \hat{\omega}_{ik-2}) \leq (\hat{\omega}_{ik} - \hat{\omega}_{ik-1}) \geq (\hat{\omega}_{ik+1} - \hat{\omega}_{ik})$, iii) chose the value at the first turning point '$\hat{\omega}_k$' as our threshold for finding the closest time series set $\hat{\mathcal{X}}$. Fig. 2a shows the DTW distance measures from $i = 1^{th}$ subject to the rest of the subjects. Fig. 2b shows the GWG trend for the closest subjects found using the proposed heuristic in the training phase (yellow region) which is similar in the test region (to be forecasted).

## III. EXPERIMENTS AND DISCUSSION

### *A. Data*

The weight data was collected from 80 women using a WiFi-connected weight scale, Withings WS30[1]. Two midwife practices in Eindhoven, The Netherlands recruited participants that were older than 18 years old and were in their gestational week 10 or earlier. The participants were asked to log their weights weekly and the recorded weight data was sent to the cloud via a mobile application. The participants provided an informed consent pre-data collection and the study was approved by Philips' Internal Ethics Committee for Biomedical Experiments (ICBE). The participants were $31 \pm 3.5$ years old (Mean$\pm$SD), and $1.69 \pm 0.07$ meters tall. Self-reported pre-pregnancy weight was $69 \pm 15$ kg resulting in an average pre-pregnancy BMI of $24 \pm 4$ kg/$m^2$. The delivery was around $277 \pm 10$ days and they gained $13.7 \pm 4.7$ kg during this time. Table II shows the data distribution of our sample dataset pre and post-pregnancy for under/within/over guidelines. Interestingly, our sample dataset's distribution is close to that in [1], which is obtained from a large population of more than a million women, with almost half of the women gaining above the recommended guidelines. This further strengthens the need for this study.

[1]https://www.withings.com/




(a) Closest time series calculated using *heuristics* on DTW distances ($i = 1^{th}$ subject)

(b) Closest subjects found using heuristics reduce variability for GPR forecasting

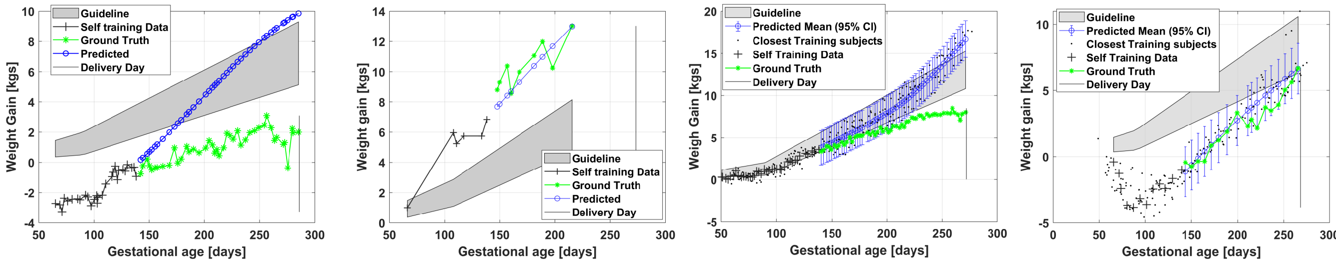
Fig. 2: DTW distance as a measure of finding dis-similarity among time series and applying heuristics to find subset from population that are closest to individual data

TABLE II: Change in distribution before and after pregnancy

| Pre-pregBMI Class | #Sub | Distribution post-pregnancy | | |
|---|---|---|---|---|
| | | Underweight | Normal | Overweight |
| Underweight | 3 | 1 | 2 | 0 |
| Normal | 45 | 11 | 15 | 19 |
| Overweight | 32 | 4 | 8 | 20 |
| **80 subjects(Class %)** | | 16(**20%**) | 25(**31.2%**) | 39(**48.8%**) |

### *B. Results*

We perform *leave-one-out* cross validation for evaluation, where training data in each iteration consists of weight gain data from N-1 population subjects and self-training data from the test subject. We normalise the absolute weight data by subtracting pre-pregnancy weight to get weight-gain data. Fig. 3 shows the progression of actual weight gain and best/worst predictions obtained using our approach. Next, we present a comparison of the two approaches among themselves and show that both outperform the state-of-the-art ARIMA approach and polynomial fitting. The performance of regression was computed using Mean Absolute Error (MAE), $MAE = \frac{1}{N}\sum_N |y(t_m^i) - y_{ref}(t_m^i)|$.




(a) Worst prediction using parametric approach, $i = 45$

(b) Best prediction using parametric approach, $i = 16$

(c) Worst prediction using non-parametric approach, $i = 55$

(d) Best prediction using non-parametric approach, $i = 40$

Fig. 3: Parametric and non-parametric approach to forecast weight gain with best and worst predictions alongside the actual weight gain data and recommended guidelines with number of training days $= 140$ for $i^{th}$ subject(s). 95% confidence interval also plotted for non-parametric approach, GPR in the predictions. (Images best viewed when zoomed 200%)

We experiment with first, second and third order polynomial based parametric approach to fit our weight-gain data. The intercept term can be omitted prior to fitting the model as the data is normalised to pass through the origin. We chose third order polynomial for parametric approach as it obtains the minimum prediction error among all other orders. The non-parametric gaussian process regression along with dynamic selection of the heuristic discussed in section II-B.1 obtains significantly accurate weight-gain predictions without fixing any parameter a-priori.

It is important to note that our parametric and non-paramteric approaches find closeness of self-training phase of the test subject to the population data in parameters or data-space respectively. Hence, our approaches implicitly assume that subjects (test and calculated population subset) similar in their training phase (i.e available weight gain data till some day) behave or show the same trend during the test phase with slight variability. Both the parametric and non-parametric approaches produce worst results shown in Fig. 3 when this assumption does not hold. This happens because the trend of the weight gain of test subject is similar to that of closest selected subset from population data during the allowed training day, but is completely different in test phase. For example, in Fig. 3c, the selected closest subset and self-training data are very close till the training day, but the test subject shows very different trend during the later stages of pregnancy. In such cases, more self-training data can improve the predictions.

### C. *Effect of number of training days*

We study the impact of availability of weight gain data for training as the pregnancy progresses and suggest a good choice of training day empirically. Fig. 4 shows that the prediction error reduces as the training data availability increases. However, in order to maximise the utility of the proposed weight-gain estimator, the training should be done as early as possible into the gestational age. We chose day '140' as the *sweet spot* because it is midway through the pregnancy where prediction capability is strong and the intervention is in time.

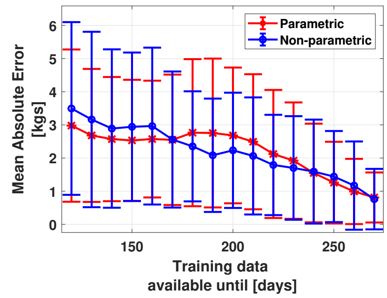



Fig. 4: Prediction error improves as more self-training data becomes available for both the approaches.

### D. *State-of-the-art*

We compare the proposed techniques with the state-of-the-art ARIMA approach for time series forecasting [6]. Forecasting using ARIMA methods require uniformly spaced samples of the time series. We introduce uniformity by linear interpolation between samples. We fit an ARIMA*(p,d,q)* model by i) enforcing equi-spaced sampling by linear interpolation, ii) performing a grid search over the hyperparameters [11] to find an optimal autoregressive order, degree of differencing, and moving average order, iii) forecasting multi-steps ahead in time to find the end-of-pregnancy gestational weight gain using the optimised hyperparameters over the training part (GWG data until day $t_d$). Fig. 5 shows that our proposed method outperforms the state-of-the-art approach in early detection and performs very closely when abundant self-training data until the start of third trimester is available. We also tested polynomial fitting approach following maximum likelihood estimation (MLE) with first to third order polynomial and order = 2 produces best results (among the orders 1 to 3). The MAE of our proposed parametric and non-parametric approach is 2.572 kgs and 2.890 kgs respectively 'mid-way' through pregnancy (day 140), whereas state-of-the-art ARIMA and MLE produce high MAE of 8.168 kgs and 16.22 kgs respectively.

### E. *Classification with respect to guidelines*

In order to further assess our method, we classify the predicted weight gain with respect to the recommended

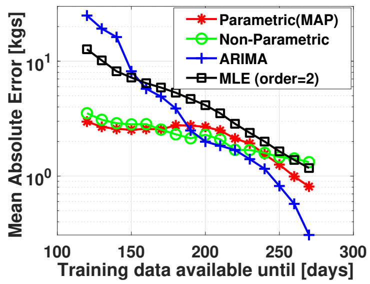



Fig. 5: Comparison of proposed approach with state-of-the-art. Proposed approach outperforms ARIMA for early intervention when training data is less and is close when abundant training data is available.

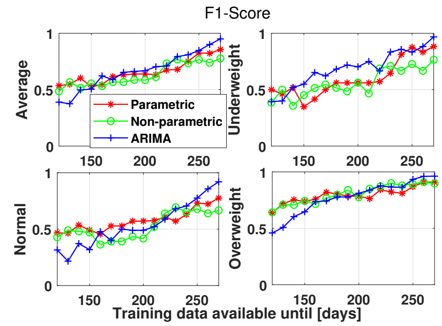



Fig. 7: Comparison of the class-wise F1-score of our approach with the state-of-the-art

guidelines into three classes, ‘underweight’, ‘normal’, ‘overweight’. In Fig. 6 and 7, we show the accuracy and F1-scores of the various methods considered. Our approach gives a maximum of 63.5% prediction accuracy at day ‘140’ for early intervention as compared to state-of-the-art, which is only 53.75% accurate. Fig. 7 also shows F1 scores for each individual class and also average F1-score across all classes.

ARIMA models the time series based on the historical data. As more and more self-training data becomes available, ARIMA based personalised models forecast better than our proposed model, which uses a-priori population data. However, in the absence of enough self-training data our proposed model out-performs the state-of-the-art approach, which is imperative for early detection of GWG trend in reducing related risks. Fig. 6 and 7 show that our approach predicts better than the state-of-the-art when training from data using 120-160 days, predicts close to state-of-the-art when data is between 161-230 days and under-performs during the last stages of the pregnancy.

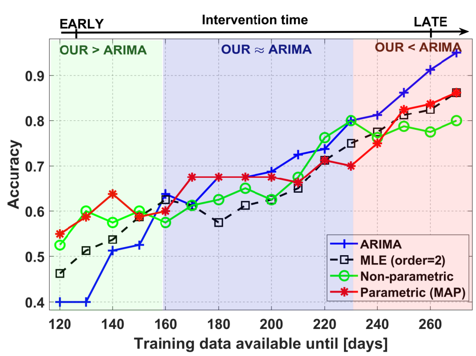



Fig. 6: Prediction accuracy of the proposed approach with respect to the state-of-the-art

## IV. CONCLUSION

We have proposed a novel excessive Gestational weight gain estimator for expecting mothers that can predict the weight gain trend in time. This can help to provide proper interventions by pre-natal care providers and to reduce risks of adverse maternal and neonatal effects of excessive or inadequate GWG. We show that our proposed algorithm outperforms the state-of-the-art approach in early intervention by utilising the power of combining a-priori information learnt from population and tuning the personal model accordingly. In the future, we would like to improve the predictions by adding additional information such as meta-data and/or fusing the two proposed approaches.

## ACKNOWLEDGMENT


This project has received funding from the European Union’s Horizon 2020 research and innovation programme under the Marie Sklodowska-Curie grant agreement No 766139.


## REFERENCES


[1] R. F. Goldstein, S. K. Abell, S. Ranasinha, M. Misso, J. A. Boyle, M. H. Black, N. Li, G. Hu, F. Corrado, L. Rode, *et al.*, “Association of gestational weight gain with maternal and infant outcomes: a systematic review and meta-analysis,” *Jama*, vol. 317, no. 21, pp. 2207–2225, 2017.

[2] R. Gaillard, B. Durmuş, A. Hofman, J. P. Mackenbach, E. A. Steegers, and V. W. Jaddoe, “Risk factors and outcomes of maternal obesity and excessive weight gain during pregnancy,” *Obesity*, vol. 21, no. 5, pp. 1046–1055, 2013.

[3] L. A. Gilmore, M. Klempel-Donchenko, and L. M. Redman, “Pregnancy as a window to future health: excessive gestational weight gain and obesity,” in *Seminars in perinatology*, vol. 39, no. 4. Elsevier, 2015, pp. 296–303.

[4] E. Oken, S. L. Rifas-Shiman, A. E. Field, A. L. Frazier, and M. W. Gillman, “Maternal gestational weight gain and offspring weight in adolescence,” *Obstetrics and gynecology*, vol. 112, no. 5, p. 999, 2008.

[5] K. M. Rasmussen, P. M. Catalano, and A. L. Yaktine, “New guidelines for weight gain during pregnancy: what obstetrician/gynecologists should know,” *Current opinion in obstetrics & gynecology*, vol. 21, no. 6, p. 521, 2009.

[6] G. E. Box, G. M. Jenkins, G. C. Reinsel, and G. M. Ljung, *Time series analysis: forecasting and control*. John Wiley & Sons, 2015.

[7] L. Clifton, D. A. Clifton, M. A. Pimentel, P. J. Watkinson, and L. Tarassenko, “Gaussian process regression in vital-sign early warning systems,” in *Engineering in Medicine and Biology Society (EMBC), 2012 Annual International Conference of the IEEE*. IEEE, 2012, pp. 6161–6164.

[8] C. E. Rasmussen, “Gaussian processes in machine learning,” in *Advanced lectures on machine learning*. Springer, 2004, pp. 63–71.

[9] D. J. Berndt and J. Clifford, “Using dynamic time warping to find patterns in time series.” in *KDD workshop*, vol. 10, no. 16. Seattle, WA, 1994, pp. 359–370.

[10] E. Keogh and C. A. Ratanamahatana, “Exact indexing of dynamic time warping,” *Knowledge and information systems*, vol. 7, no. 3, pp. 358–386, 2005.

[11] R. Shibata, “Selection of the order of an autoregressive model by akaike’s information criterion,” *Biometrika*, vol. 63, no. 1, pp. 117–126, 1976.