Learning models for writing better doctor prescriptions *

Tingting Xu[1] and Ioannis Ch. Paschalidis[2], *Fellow, IEEE*

***Abstract*— We develop a data-driven approach for learning and improving the prescription policy physicians use to treat Type 2 diabetes. Our model combines regression, classification and strategy optimization. We use regression algorithms to predict the outcomes of prescriptions, and then adopt a parameterized classification method to learn the physicians' prescription policy. Finally, we improve the prescription policy by optimizing over the parameters in the prescription policy model. Compared with the original prescription policy, patients who shift their treatment according to the recommended policy see significant blood glucose reduction on average. The proposed prescription recommendations offer a better therapeutic effect than the state-of-art deterministic algorithms. Our framework can also be applied to improving the prescription policy for other diseases.**

I. INTRODUCTION

There is an increasing percentage of the worldwide population suffering from chronic diseases, mainly due to lifestyle and diet choices. Diabetes, in particular, is one of the leading chronic diseases and about 30.8 billion was spent in the U.S. in 2017 on anti-hypoglycemic medications. Type 2 diabetes patients account for $90\% - 95\%$ among all patients with diabetes [1]. Improving the treatment efficiency is important to improve patients' long-term life quality, as well as reduce health care costs. A promising solution is to provide personalized therapy recommendations for heterogeneous patients based on their health history and characteristics [2].

Some existing research has modeled the prescription recommendation problem using ideas from dynamic systems and control theory. For instance, [3] proposed a stress intervention mechanism based on fuzzy control theory. [4] used a Markov Decision Process (MDP) framework to develop intervention strategies for breast and ovarian cancer. [5] used Reinforcement Learning (RL) to learn dynamics and cost functions.

Most of the related existing work has a number of key limitations: (a) Unrealistic assumptions: RL-based methods require simulation or accurate models to capture the (elusive) interaction between patients and treatments. (b) Lack of generalization: the design of a dynamic system is valid for a specific disease and the models may vary from disease to disease. (c) Non-personalized recommendations: MDP-based models assign patients into different "states" and provide uniform recommendations at each state. (d) Limited scale: MDPs suffer from the well known curse of dimensionality.

Our work in this paper develops a new "data-driven" framework for personalized treatment recommendations that does not require simulation. One of the major challenges of personalized therapy recommendations is to assess patient-specific effects of treatments, because the effects of the same drugs may vary from patient to patient. Although clinical trials can be utilized to evaluate the effects of one treatment, it is impossible to assess the counterfactual efficacy of alternative treatments. Besides, large-scale clinical trials are economically ineffective and time consuming. Fortunately, the increasing availability of *Electronic Health Records (EHRs)* gives us access to large amounts of patients' medical data, from which we can extract information about physicians' prescription policies and patient-specific treatment efficacy.

There are two natural strategies for personalized treatment recommendations based on data. The first strategy is to directly learn from data a policy which accurately predicts physicians' prescriptions; this method utilizes doctors' domain knowledge and the recommendations will be more easily acceptable by doctors; however, it is difficult to improve the physicians' policy by only learning from them. Another recommendation strategy is to directly select the treatment which optimizes the *predicted* prescription effects [6]; this approach may produce problematic recommendations especially when the predictive model has low accuracy.

Our work combines these two strategies. Specifically, we first learn the patient-specific effects of treatments using regression. Then, we design a model to accurately predict a doctor's prescription policy based on patients' EHRs. Finally, we design a method to improve the current prescription policies based on the prediction model of treatment effects and the estimated doctors' prescription policy model. Rather than only using the prescription effect regression models as in [6], we also utilize information from the learned prescription models. This reduces the impact of low-precision regression models and results in better treatment effects in real experiments. Furthermore, our model produces a probability as a confidence score for each recommendation, which provides further information physicians can use.

The remainder of this article is organized as follows. Sec. II decomposes the problem into three sub-problems: treatment effect regression, learning doctors' prescription policies, and improving these policies. The detailed methods for these subproblems are described in Secs. III–V. Sec. VI illustrates the effectiveness of our approaches using a real

* Research partially supported by the NSF under grants DMS-1664644, CNS-1645681, IIS-1237022, and CCF-1527292, by the ONR under MURI grant N00014-16-1-2832, by the NIH under grant 1UL1TR001430 to the Clinical & Translational Science Institute at Boston University, and by the Boston University Digital Health Initiative.

[1] Center for Information and Systems Eng., Division of Systems Eng., Boston University, e-mail: tingxu@bu.edu.

[2] Department of Electrical and Computer Eng., Division of Systems Eng., and Dept. of Biomedical Eng., Boston University, 8 St. Mary's St., Boston, MA 02215, e-mail: yannisp@bu.edu, http://sites.bu.edu/paschalidis/.

EHR dataset. Sec VII draws some conclusions.

II. The Problem

As alluded to earlier, the proposed approach consists of three components: (a) developing a predictive model to estimate the prescription effects; (b) learning the doctors' prescription policy based on patients' EHRs; and (c) improving the prescription policy to induce better outcomes.

For subproblem (a), we will use multiple regression models to assess the effect of various treatments, based on patients' EHR. If the predicted effects from multiple regression models are consistent, we can be highly confident in the predictions. For subproblem (b), we consider learning the doctors' prescription policy as a multi-class classification problem and propose a parametric model, where the input features are the predicted prescription results and the patient's medical history, and the output is the prescription policy. For subproblem (c), by appropriately modifying the parametric policy, we are able to improve the current prescription policy and achieve better glycemic control.

Notation: We use bold lowercase and uppercase letters for vectors and matrices, respectively. Prime denotes transpose. All vectors are column vectors. $\|\cdot\|_1$ and $\|\cdot\|_2$ denote the ℓ_1 and ℓ_2 norms, respectively. The matrix $\mathbf{X} \in \mathbb{R}^{n\times d}$ denotes patients' records, with rows corresponding to patients and columns to patient features. The vector $\mathbf{y} \in \mathbb{R}^n$ contains the prescription effect (future glucose) of the physician's prescribed treatment for the n patients. Assume there are L types of treatments. We split all the records based on the treatment type and denote by $\mathbf{X}^\ell \in \mathbb{R}^{n_\ell \times d}$ the n_ℓ records under treatment ℓ, and by $\mathbf{y}^\ell \in \mathbb{R}^{n_\ell}$ the true effect of that treatment, for $\ell = 1, \ldots, L$.

III. Predictive model of the treatment effects

We use several classic regression models to predict the effect of the ℓ-th treatment type for all patients. A simple linear model is the *Least Absolute Shrinkage and Selection Operator (LASSO)* regression [7]:

$$(\hat{\mathbf{w}}_{LA}^\ell, \hat{b}_{LA}^\ell) = \arg\min_{\mathbf{w}^\ell, b^\ell} \|\mathbf{X}^\ell \mathbf{w}^\ell + b^\ell - \mathbf{y}^\ell\|_2^2 + \lambda \|\mathbf{w}^\ell\|_1, \quad (1)$$

where $\mathbf{w}^\ell$ contains the regression coefficients, b^ℓ is the offset, and the parameter λ trades off training loss and model sparsity. In practice, λ is determined by cross validation. The Ordinary Least Square (OLS) regression is a special case when $\lambda = 0$. Problem (1) is convex and easily solved for large instances.

Non-parametric and nonlinear regression models, like *Random Forests (RF)* and *k-nearest-neighbor (kNN)* based methods, are also considered. RF regression is an ensemble algorithm to build a complex model that trains multiple decorrelated decision trees and forms average predictions from the tree models [8]. Given a new sample, the kNN model uses the average of its k nearest neighbors in the training set as the prediction. The model lacks robustness to noise due to the intrinsic local information property, and its performance is known to be affected by computational complexity in high-dimensional cases [9].

We will also consider a *weighted kNN (WkNN)* algorithm, a variant of kNN used in [6], to provide a comprehensive comparison. Instead of using the Euclidean distance as in kNN, the WkNN model uses a weighted Euclidean distance, where the weights of features come from the coefficients of a linear regression model. We use the coefficients from LASSO as weights in our implementation, because it outperforms OLS regression in terms of regression accuracy.

We split the data into a training and test set, and use 5-fold cross-validation within the training set to determine the model hyper-parameters. With the training set, we can train the regression model using any of the aforementioned regression algorithms. For any sample $\mathbf{x}$, we denote by $\hat{y}(\mathbf{x}, \ell)$ the predicted effect (future glucose) under treatment type ℓ. The test set can be used to evaluate the accuracy of regression models; a classic performance metric is the coefficient of determination:

$$R^2 = 1 - \frac{\sum_i (\hat{y}_i - y_i)^2}{\sum_i (\frac{1}{n}\sum_i y_i - y_i)^2},$$

which describes the proportion of the variance in the target variable explained by the regression algorithm.

IV. Learning the physicians' prescription policy

This section aims to learn the physicians' treatment decision policy. To that end, we will use the patient's features and the predicted patient-specific effects of various treatments obtained in Sec. III.

Recall that $\mathbf{X} \in R^{n\times d}$ contains the patient records and let $\mathbf{u} \in \{1, \cdots, L\}^n$ be the doctors' prescriptions for each patient. We let $\hat{\mathbf{Y}} \in \mathbb{R}^{n\times LM}$ represent the treatment effects predicted by all M regression methods $\{\mathcal{A}_1, \ldots, \mathcal{A}_M\}$, where the $(m-1)L+l$ column of $\hat{\mathbf{Y}}$ represents the predicted effects under prescription type ℓ from the mth algorithm $\mathcal{A}_m$. We denote by $\mathbf{Z} = [\mathbf{X}\ \hat{\mathbf{Y}}]$ the combined input features, and split them into a training set $\mathbf{Z}_{train}$ and a test set $\mathbf{Z}_{test}$.

Specifically, our objective is to obtain a mapping from the input features $\mathbf{Z}$ to actions $\mathbf{u}$. This is a multi-class classification problem and can be solved by the multinomial logistic regression (also called softmax regression) algorithm [10]. The probability of prescribing a type-ℓ treatment for the sample $\mathbf{z}_i = (\mathbf{x}_i, \hat{\mathbf{y}}_i)$ is:

$$P(u_i = \ell | \mathbf{z}_i) = \frac{\exp((\boldsymbol{\beta}^\ell)' \boldsymbol{\phi}(\mathbf{z}_i))}{\sum_{r=1}^L \exp((\boldsymbol{\beta}^r)' \boldsymbol{\phi}(\mathbf{z}_i))}, \quad (2)$$

where $\boldsymbol{\beta}^\ell$ is a weight vector and $\boldsymbol{\phi}(\mathbf{z}_i)$ is a feature vector transforming the ith record $\mathbf{z}_i$; the latter can be obtained using feature selection methods such as recursive feature selection [11].

Using training data $\mathbf{Z}_{train}$ for n_{train} patients and corresponding output actions $\mathbf{u}_{train}$, we obtain model parameters $\boldsymbol{\beta}^\ell$ in (2) by solving an ℓ_1-regularized multinomial logistic regression problem:

$$\begin{aligned} \min_{\boldsymbol{\beta}^\ell} \quad & -\sum_{i=1}^{n_{train}} \log \frac{\exp((\boldsymbol{\beta}^{u_i})' \boldsymbol{\phi}(\mathbf{z}_i))}{\sum_{\ell=1}^L \exp((\boldsymbol{\beta}^\ell)' \boldsymbol{\phi}(\mathbf{z}_i))} \\ \text{s.t.} \quad & \|\boldsymbol{\beta}^\ell\|_1 \le \gamma, \quad \forall \ell. \end{aligned} \quad (3)$$

This is a convex optimization problem and can be solved efficiently, e.g., by an incremental gradient method with a linear convergence rate [12].

In the testing phase, for a patient $\mathbf{x}_{test}$, we first use the regression approaches in Sec. III to predict the treatment effects $\hat{\mathbf{y}}_{test}$, and then, given $\mathbf{z}_{test} = (\mathbf{x}_{test}, \hat{\mathbf{y}}_{test})$, we predict the most likely treatment:

$$u^*(\mathbf{z}_{test}) = \arg\max_{\ell} \frac{\exp((\hat{\boldsymbol{\beta}}^{\ell})'\boldsymbol{\phi}(\mathbf{z}_{test}))}{\sum_{r=1}^{L}\exp((\hat{\boldsymbol{\beta}}^{r})'\boldsymbol{\phi}(\mathbf{z}_{test}))},$$

where $\hat{\boldsymbol{\beta}}^{\ell}$ is the optimal parameter derived from (3). The out-of-sample classification accuracy is used to evaluate the model accuracy and is defined as the percentage of correctly classified samples among all samples in the test set.

V. Improving the current prescription policy

This section seeks to derive a recommended prescription policy with better treatment effects (lower glucose for diabetic patients) by optimizing over the parameters of the learned current policy. Given a sample $\mathbf{x}$, a natural strategy is to select an action which optimizes the predicted effects, i.e.,

$$u_A^*(\mathbf{x}) = \arg\min_{u} \hat{y}_A(\mathbf{x}, u),$$

using the predictive model $\hat{y}_A(\mathbf{x}, u)$ of each regression algorithm $A \in \{\mathcal{A}_1, \cdots, \mathcal{A}_M\}$.

At the same time, we prefer not to deviate too far from the current prescription policy, because treatment changes may be costly and introduce transient effects which may be undesirable. An additional reason for "constraining" prescriptions is that we can only rely on the predictive models to predict the effect of the prescription and these models are not always extremely accurate.

In this context, we formulate a new constrained softmax regression problem, where the input features are still $\boldsymbol{\phi}(\mathbf{z}_i)$ but the target labels are $u_A^*(\mathbf{x}_i)$, forcing the model to learn actions which optimize the predicted treatment effects. We solve

$$\begin{aligned} \min_{\boldsymbol{\beta}^{\ell}} \quad & -\sum_{i=1}^{n_{train}} \log \frac{\exp((\boldsymbol{\beta}^{u_A^*(\mathbf{x}_i)})'\boldsymbol{\phi}(\mathbf{z}_i))}{\sum_{\ell=1}^{L}\exp((\boldsymbol{\beta}^{\ell})'\boldsymbol{\phi}(\mathbf{z}_i))} \\ \text{s.t.} \quad & \|\boldsymbol{\beta}^{\ell}\|_1 \le \gamma, \quad \forall \ell. \\ & \|\hat{\boldsymbol{\beta}}^{\ell} - \boldsymbol{\beta}^{\ell}\|_2 < \eta, \quad \forall \ell, \end{aligned} \tag{4}$$

where γ and η are fixed positive numbers, and $\hat{\boldsymbol{\beta}}^{\ell}, \ell = 1, \ldots, L$ are optimal parameters obtained from (3). Notice that the last constraint, restricts the difference between the parameters $\boldsymbol{\beta}^{\ell}$ of the new policy and the corresponding ones for the original prescription policy.

Obtaining an optimal solution $\boldsymbol{\beta}^{\ell}$ to (4) defines a probabilistic recommended prescription. Specifically, for a patient with features $\mathbf{z}$ one can select treatment l with probability

$$P(\hat{u}(\mathbf{z}) = \ell) = \frac{\exp((\boldsymbol{\beta}^{\ell})'\boldsymbol{\phi}(\mathbf{z}))}{\sum_{r=1}^{L}\exp((\boldsymbol{\beta}^{r})'\boldsymbol{\phi}(\mathbf{z}))}.$$

These probabilities provide additional information to physicians, essentially stating the level of preference (or confidence) in specific actions [13]. To obtain a deterministic policy, one can simply recommend the most likely proscription

$$u^*(\mathbf{z}) = \arg\max_{\ell} \frac{\exp((\boldsymbol{\beta}^{\ell})'\boldsymbol{\phi}(\mathbf{z}))}{\sum_{r=1}^{L}\exp((\boldsymbol{\beta}^{r})'\boldsymbol{\phi}(\mathbf{z}))}. \tag{5}$$

When evaluating effects of the recommended prescription policy, we will use this deterministic prescription policy and the regression algorithms in Sec. III to predict its effects. The treatment effects will be compared with the corresponding effects of the original physicians' prescription policy.

VI. Experimental Results

In this section, we apply our framework to a data set with Type 2 diabetic patients. The data set is from the Boston Medical Center (BMC) – the largest safety-net hospital in New England.

A. Data descriptions and preprocessing

We extract all demographic and medical records during 01/01/2001-12/31/2007 for patients who meet the following criteria: (a) received at least one antiglycemic (blood glucose regulation) agent prescription, including insulin, metformin, sulfonylureas, etc; (b) did not have a diagnosis of Type 1 diabetes; and (c) had at least two laboratory test records of long-term blood glucose, measured by glycated hemoglobin (HbA1c) which reflects average glycemia over a 3-month period. The HbA1c values represent the percentage of glycated hemoglobin among total hemoglobin.

Our dataset includes 4116 such patients in total, for which we capture: (1) Demographics: gender, age, race; (2) Diagnoses: hypertensive disease, skin infections, etc. (3) Procedures: e.g., transfusion of packed cells; (4) Admissions: e.g., diabetes, heart failure, chest pain; (5) Service Department: inpatient, emergency room, etc.; (6) Lab Tests: hematology, chemistry, urinalysis tests; (7) Vital Signs: e.g., blood pressure, temperature, pulse; (8) Blood Glucose Regulation Agents: insulin, anti-hypoglycemic, oral hypoglycemic agents, etc.

We conduct a data preprocessing procedure to organize all patients' medical information in a uniform way. Specifically: (a) For each patient, we group all patients' EHR history into 3-month windows starting from one month before their first HbA1c measurement. The selection of period length is based on the 3-month observation period adopted in the diabetes standards-of-care guidelines [14].
(b) To summarize the information of a patient in each 3-month window, the mean of continuous variables and occurrence counts of categorical features are calculated, respectively.
(c) There are mainly two types of glucose regulation drugs to treat Type 2 diabetes, oral-type drugs including metformin, sulfonylureas, etc.; and more powerful injectable-type drugs like insulin, GLP-1 receptor agonists, etc. [14]. According to the drug types used in a 3-month period, we group all

treatments into 3 types: no treatment, treatment with only oral drugs, and treatment with injectable drugs involved.

(d) We are interested in predictions of key variables (HbA1c) and treatment recommendations during a patient's 3-month "target" period, which can be any of the 3-month periods available for this patient. We will use as input to predictive/prescriptive models, patient features over the 2 preceding 3-month periods. Specifically, using index t to denote the target period, a typical input sample is $\mathbf{x}_t = (\mathbf{v}_{t-2}, \mathbf{v}_{t-1}, \mathbf{v}_t)$, where $\mathbf{v}_{t-i}$ contains EHR records in the period which is i periods before the target period. Note that patients with long medical history can have multiple such samples by shifting the target periods. We denote by u_t the current physicians' prescription during the target-period, and by y_{t+1} the true effect of physicians' treatment or average HbA1c in the next period.

(e) We remove samples when there are no HbA1c values during the target-period t or the next-period $t+1$. For other missing values, we replace the missing continuous valued ones with their median values, and use an additional indicator variable to represent missing categorical variables.

These preprocessing procedures lead to $15,177$ final sample records.

B. Experimental Results

Samples are split into a training set and a test set, with a percentage of $2/3$ and $1/3$, respectively. For the samples in the training set, we assume that the current-period (target-period) treatment types and the true future (next-period) HbA1c values are available. The whole learning and prescription recommendation process is repeated for 10 runs, with randomly selected training sets.

1) Modeling of the prescription effects: In each run, we group all samples in the training set according to the treatment types during the current period. Based on all training samples with treatment type ℓ, a treatment effect prediction model is trained by the regression algorithms in Sec. III. The inputs for models are $\{\mathbf{X}^\ell, y^\ell\}$, in which $\mathbf{X}^\ell \in R^{n_l \times d}$ is the set of n_l records where the patients' current treatment type is ℓ, and y^ℓ is the corresponding treatment effect (future HbA1c), for all $\ell = 1, 2, 3$. We repeat the treatment regression experiments for 10 runs with randomly selected training sets. Table I shows the mean and standard deviation (std) of the out-of-sample accuracy (R^2) of all regression algorithms under various treatment types. Random Forests (RF) achieve the highest accuracy and kNN the lowest. We will use the predicted treatment effects from all regression models except kNN in learning and improving the doctors' prescriptions.

Exploiting the feature selection capability of LASSO, we show in Table II the ten most important features for predicting future HbA1c under oral medication; the table lists results from the model with the highest accuracy through 10 runs. The features are ranked by the absolute values of their corresponding coefficients in the model. Larger coefficients imply a stronger positive association between the feature and the response. The most important features include: HbA1c history and oral treatment history, various types of glucose, blood pressure, age and mean corpuscular hemoglobin, which is the average mass of hemoglobin in red blood cells. Endocrinologists would agree that there is a high correlation between these factors and HbA1c.

TABLE I

R^2 OF REGRESSION METHODS FOR FUTURE HBA1C PREDICTION

		$\ell = 1$	$\ell = 2$	$\ell = 3$
LASSO	mean(R^2)	0.55	0.51	0.45
	std(R^2)	0.01	0.01	0.02
WkNN	mean(R^2)	0.55	0.52	0.47
	std(R^2)	0.01	0.01	0.02
kNN	mean(R^2)	0.28	0.23	0.25
	std(R^2)	0.01	0.01	0.002
RF	mean(R^2)	0.57	0.53	0.51
	std(R^2)	0.01	0.01	0.01

TABLE II

IMPORTANT FEATURES FROM LASSO FOR PREDICTING FUTURE HBA1C UNDER ORAL MEDICATION

Rank	Feature	Coef
1	Target-period HbA1c	0.67
2	HbA1c of the period 6 months before the target period	0.18
3	HbA1c of the period before the target period	0.16
4	Target-period point-of-care glucose	0.14
5	Target-period blood glucose tested by finger stick	0.13
6	Target-period erythrocyte mean corpuscular hemoglobin (MCH RBC Qn)	-0.08
7	Hospital admission due to deep vein thrombophlebitis with a comorbidity	-0.08
8	Oral treatment was prescribed during the period before the target period	0.07
9	Target-period blood pressure	0.07
10	Age	0.07

2) Learning the current prescription policy: For each sample $\mathbf{x}$ we first obtain the predicted HbA1c under each possible treatment using the models of Sec. VI-B.1 (except kNN); let $\hat{\mathbf{y}}(\mathbf{x})$ denote the vector with these predictions. We form a new feature vector $\mathbf{z}(\mathbf{x}) = (\mathbf{x}, \hat{\mathbf{y}}(\mathbf{x}))$ and, as discussed in Sec. IV, transform it to $\boldsymbol{\phi}(\mathbf{z})$ to provide an input to (2) from which we predict the current doctor's prescription by selecting the most likely one.

Table III lists the mean and standard deviation (std) of the accuracy, defined as the percentage of correctly predicted prescriptions. Rows corresponding to $\hat{y}_{LA}$, $\hat{y}_{WkNN}$, and $\hat{y}_{RF}$ used the predicted future HbA1c from LASSO, WkNN, and RF regression, respectively, as features $\boldsymbol{\phi}(\mathbf{z})$ in order to predict the current prescription. The last row uses features from the patient's EHR and demographic features. It is not surprising that using more information (as the last row does) leads to significantly better accuracy. It is also interesting that the best method for predicting future HbA1c (RF) does not provide the best input feature for predicting the current prescription.

To understand the commonality of features used for predicting the response and the treatment, we extracted the

30 most important features from the best LASSO regression model for each treatment type obtained in Sec. VI-B.1. Removing duplicate features yields 71 features which are informative in predicting future HbA1c. Similarly, we considered the most important features used by the models developed in this section to learn the doctor's prescription policy, which yielded 53 features. These two distinct sets of features have only 14 features in common. This implies that instead of considering only the predicted HbA1c (as in [6]) to make a treatment recommendation, physicians' prescription decisions seem to rely on additional information from the patient's medical history. This motivates the type of models we consider in the next section, seeking to make prescription decisions using both predicted future HbA1c and additional information from the EHR.

TABLE III
THE ACCURACY OF LEARNING PRESCRIPTIONS

Features	Accuracy (Mean)	Accuracy (STD)
$\hat{y}_{LA}$	0.60	0.016
$\hat{y}_{WkNN}$	0.51	0.004
$\hat{y}_{RF}$	0.49	0.004
EHR and predicted future HbA1c	0.83	0.003

3) Improving the current prescription policy: The parametric softmax regression model of Sec. V is adopted to improve the current prescription policy. We compare our recommendation strategy, called rec_2, with the recommendations from [6], called rec_1:

rec_1 (*policy in [6]*): For each sample $\mathbf{x}_i$, the recommended treatment for the target period is $u^{rec_1}(\mathbf{x}_i) = \arg\min_u \hat{y}_{t+1}(\mathbf{x}_i, u)$ only if the predicted future effect $\hat{y}_{t+1}(\mathbf{x}_i, u^{rec_1}(\mathbf{x}_i))$ is significantly better than the current HbA1c $y_t(\mathbf{x}_i, u_{t-1}(\mathbf{x}_i))$, otherwise the treatment prescribed in the previous period $u_{t-1}(\mathbf{x}_i)$ is adopted. Specifically,

$$u_t^{rec_1}(\mathbf{x}_i) = \begin{cases} u_{t-1}(\mathbf{x}_i), \text{ if } \min_u \hat{y}_{t+1}(\mathbf{x}_i, u) > \eta\ y_t(\mathbf{x}_i, u_{t-1}(\mathbf{x}_i)), \\ \arg\min_u \hat{y}_{t+1}(\mathbf{x}_i, u), \text{ otherwise.} \end{cases}$$

The threshold η represents the level of conservatism which can be set by the decision makers; we use $\eta = 95\%$ in our experiments. Other than the patients' current HbA1c and the previous-period treatment type, this recommendation strategy only relies on the predicted future HbA1c.

rec_2 (*our policy*): As described in Sec. V, for each patient with features $\mathbf{x}_i$ in the training set, the true HbA1c at period $t+1$, $y_{t+1}(\mathbf{x}_i, u_t(\mathbf{x}_i))$, and the current prescription $u_t(\mathbf{x}_i)$ are both available. We first create new treatment labels which optimize the predicted prescription effects:

$$u_t^*(\mathbf{x}_i) = \begin{cases} u_t(\mathbf{x}_i), \text{ if } \min_u \hat{y}_{t+1}(\mathbf{x}_i, u) > \eta\ y_{t+1}(\mathbf{x}_i, u_t(\mathbf{x}_i)), \\ \arg\min_u \hat{y}_{t+1}(\mathbf{x}_i, u), \text{ otherwise.} \end{cases}$$

This rule adopts the labels $u_t^*(\mathbf{x}_i) = \arg\min_u \hat{y}_{t+1}(\mathbf{x}_i, u)$ if the predicted resulting HbA1c $\hat{y}_{t+1}(\mathbf{x}_i, u^*(\mathbf{x}_i))$ leads to a significant improvement over the true future HbA1c, otherwise the current-period prescribed treatment is being used. Then, a softmax regression model is trained according to equation (4) with patients' input features $\mathbf{z}_i$ and new labels $u_t^*(\mathbf{x}_i)$, The recommended treatment for sample $\mathbf{z}_i$ is the predicted treatment label obtained from (5):

$$u^{rec_2}(\mathbf{x}_i) = u^*(\mathbf{z}_i).$$

Both of the above recommended policies depend on the predicted prescription effects, hence, each predictive model results in two corresponding policy recommendations (Tables V – VII). We compare the future predicted HbA1c resulting from the two prescription recommendations on the patients whose recommended treatments shift (vary) from their *current-period* original treatments prescribed by physicians. Table IV shows the average counts and ratios of shifted treatments under the two prescription policy recommendations (columns) based on various regression models (rows).

Tables V to VII present the *relative* HbA1c improvement for patients whose treatment is modified. In each of these tables: (a) the 2nd column shows the average true future HbA1c for patients who use the original physicians' prescription; (b) the 3rd to 5th columns list the average future HbA1c predicted by different regression models when using the recommended prescriptions, and the last column shows the average of these predictions; (c) the "Improvement" rows show the *relative* improvement of each recommendation compared to the original prescription policy. [1] For instance, according to the LASSO-based recommendations in Table V, rec_1 leads to a 2.3% average relative improvement from the original future HbA1c of 8.24% to 8.06%, whereas rec_2 leads to a 7.0% average relative improvement from the original future HbA1c of 8.35% to 7.77%. In all Tables V-VII, our recommendation strategy rec_2 consistently outperforms rec_1.

TABLE IV
COUNTS AND RATIOS OF PATIENTS WITH MODIFIED TREATMENT

	Shifted Treatments	rec_1	rec_2 (ours)
$u_{\mathbf{LASSO}}$	Count	5697	5207
	(Ratio)	(37.54%)	(34.31%)
$u_{\mathbf{WKNN}}$	Count	7959	7161
	(Ratio)	(52.44%)	(47.18%)
$u_{\mathbf{RF}}$	Count	3795	4902
	(Ratio)	(25.00%)	(32.30%)

4) Analysis of the recommended prescriptions: To understand the patients' treatment distributions under the original and the recommended prescription policies, we analyze an illustrative example when using the LASSO-based treatment recommendations from rec_2 with maximal relative improvement through 10 runs. All patients' records are grouped by

[1] All HbA1c values are percentages but we omit the % symbol for brevity.

TABLE V

RELATIVE HBA1C IMPROVEMENT FOR PATIENTS WHOSE TREATMENT HAS BEEN MODIFIED BASED ON LASSO REGRESSION

$u_{\mathbf{LASSO}}$	Original	$\bar{\mathbf{y}}_{LA}$	$\bar{\mathbf{y}}_{WkNN}$	$\bar{\mathbf{y}}_{RF}$	Avg($\bar{\mathbf{y}}$)
rec_1	8.24	7.85	8.21	8.10	8.06
(Improvement)	(0)	(4.7%)	(0.4%)	(1.7%)	(2.3%)
rec_2	8.35	7.66	7.88	7.77	7.77
(Improvement)	(0)	(8.3%)	(5.7%)	(7.0%)	(7.0%)

TABLE VI

RELATIVE HBA1C IMPROVEMENTS FOR PATIENTS WHOSE TREATMENT HAS BEEN MODIFIED BASED ON WEIGHTED kNN REGRESSION

$u_{\mathbf{WkNN}}$	Original	$\bar{\mathbf{y}}_{LA}$	$\bar{\mathbf{y}}_{WkNN}$	$\bar{\mathbf{y}}_{RF}$	Avg($\bar{\mathbf{y}}$)
rec_1	7.57	7.53	7.05	7.59	7.39
(Improvement)	(0)	(0.5%)	(6.9%)	(-0.3%)	(2.4%)
rec_2	7.90	7.61	7.20	7.58	7.47
(Improvement)	(0)	(3.6%)	(8.7%)	(4.0%)	(5.4%)

their current-period HbA1c values using severity cutpoints $[5.7\%, 6.5\%, 7.0\%, 9.0\%]$ [14]. We plot the counts of patients' records versus their HbA1c group labels in Fig. 1. For the first group (HbA1c $\leq 5.7\%$), our recommendations lead to less patients using drugs. For the 2nd–4th group of patients (HbA1c $\in [6.5\%, 9.0\%]$), the number of patients who are given oral drugs increases while the number of patients who do not take drugs or are given injectable drugs decreases. For the 5th group (HbA1c $\geq 9.0\%$), the number of patients using oral and injectable agents increases, while the number of patients who do not take drugs decreases. This is consistent with the diabetes care guidelines [14]. With our recommendations, prescriptions are more in line with standard-of-care guidelines and HbA1c is better controlled.

VII. CONCLUSIONS

We proposed a framework for learning and improving doctor prescriptions. Our approach combines procedures for predicting the effect of various treatments, learning from data a policy used by physicians, and optimizing the prescription policy to improve outcomes. We developed a multi-class classification model which was able to learn the physicians' prescription policy reflected in the data. By incorporating information from the predicted outcome of the treatment and the learned current prescription policy, we proposed an approach to improve the current prescription policy, reducing the influence of low-accuracy regression models and leading to better patient outcomes. We applied our methods to the problem of selecting a drug class to treat Type II diabetes. The proposed framework can also be adapted to making personalized treatment recommendations for other diseases.

TABLE VII

RELATIVE HBA1C IMPROVEMENTS FOR PATIENTS WHOSE TREATMENT HAS BEEN MODIFIED BASED ON RANDOM FORESTS REGRESSION

$u_{\mathbf{RF}}$	Original	$\bar{\mathbf{y}}_{LA}$	$\bar{\mathbf{y}}_{WkNN}$	$\bar{\mathbf{y}}_{RF}$	Avg($\bar{\mathbf{y}}$)
rec1	8.67	8.14	8.43	8.10	8.223
(Improvement)	(0.00%)	(6.15%)	(2.80%)	(6.52%)	(5.16%)
rec2	8.55	7.85	8.01	7.79	7.881
(Improvement)	(0.00%)	(8.17%)	(6.30%)	(8.86%)	(7.77%)

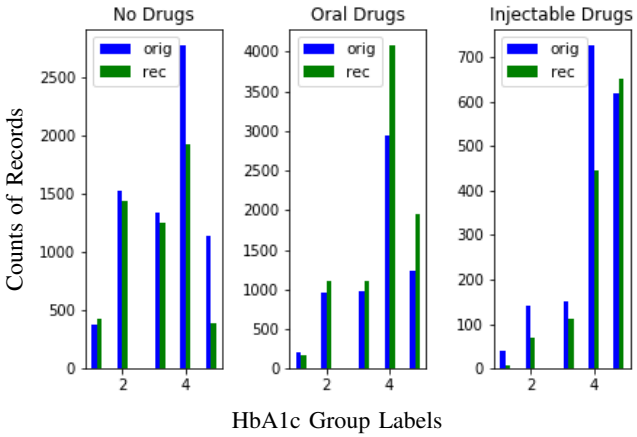

Fig. 1. Distribution of original and recommended prescriptions for various HbA1c subgroups.

REFERENCES

[1] American Diabetes Association, "Economic costs of diabetes in the US in 2017," *Diabetes Care*, vol. 41, no. 5, pp. 917–928, 2018.

[2] S. Mathur and J. Sutton, "Personalized medicine could transform healthcare," *Biomedical reports*, vol. 7, no. 1, pp. 3–5, 2017.

[3] J. M. Calderón and L. G. Jaimes, "A fuzzy control-based approach for the selection of health interventions," in *Consumer Communications & Networking Conference (CCNC), 2018 15th IEEE Annual*. IEEE, 2018, pp. 1–6.

[4] M. Abdollahian and T. K. Das, "A MDP model for breast and ovarian cancer intervention strategies for BRCA1/2 mutation carriers," *IEEE journal of biomedical and health informatics*, vol. 19, no. 2, pp. 720–727, 2015.

[5] H. Asoh, M. S. S. Akaho, T. Kamishima, K. Hasida, E. Aramaki, and T. Kohro, "An application of inverse reinforcement learning to medical records of diabetes treatment," in *ECMLPKDD2013 Workshop on Reinforcement Learning with Generalized Feedback*, 2013.

[6] D. Bertsimas, N. Kallus, A. M. Weinstein, and Y. D. Zhuo, "Personalized diabetes management using electronic medical records," *Diabetes care*, vol. 40, no. 2, pp. 210–217, 2017.

[7] R. Tibshirani, "Regression shrinkage and selection via the lasso," *Journal of the Royal Statistical Society. Series B (Methodological)*, pp. 267–288, 1996.

[8] L. Breiman, "Random forests," *Machine learning*, vol. 45, no. 1, pp. 5–32, 2001.

[9] A. Hinneburg, C. C. Aggarwal, and D. A. Keim, "What is the nearest neighbor in high dimensional spaces?" in *26th Internat. Conference on Very Large Databases*, 2000, pp. 506–515.

[10] D. W. Hosmer Jr, S. Lemeshow, and R. X. Sturdivant, *Applied logistic regression*. John Wiley & Sons, 2013, vol. 398.

[11] I. Guyon, J. Weston, S. Barnhill, and V. Vapnik, "Gene selection for cancer classification using support vector machines," *Machine learning*, vol. 46, no. 1-3, pp. 389–422, 2002.

[12] A. Defazio, F. Bach, and S. Lacoste-Julien, "SAGA: A fast incremental gradient method with support for non-strongly convex composite objectives," in *Advances in neural information processing systems*, 2014, pp. 1646–1654.

[13] A. Chakrabarty, F. J. Doyle, and E. Dassau, "Deep learning assisted macronutrient estimation for feedforward-feedback control in artificial pancreas systems," in *2018 Annual American Control Conference (ACC)*. IEEE, 2018, pp. 3564–3570.

[14] American Diabetes Association, "8. pharmacologic approaches to glycemic treatment: Standards of medical care in diabetes—2018," *Diabetes Care*, vol. 41, no. Supplement 1, pp. S73–S85, 2018.