Hybrid Predictive Modelling for Finding Optimal Multipurpose Multicomponent Therapy

Vladislav V. Pavlovskii [1[0000-0002-1690-9812]], Ilia V. Derevitskii[1[0000-0002-8624-5046]], Sergey V. Kovalchuk [1[0000-0001-8828-4615]]

[1]ITMO University, Saint Petersburg, Russia
ivderevitckii@itmo.ru

Abstract. This study presents a new hybrid approach to predictive modelling of disease dynamics for finding optimal therapy. We use existing methods, such as expert-based modelling methods, models of system dynamics and ML methods in compositions together with our proposed modelling methods for simulating treatment process and predicting treatment outcomes depending on the different therapy types. Treatment outcomes include a set of treatment-goal values, therapy types include a combination of drugs and treatment procedures. Personal therapy recommendation by this approach is optimal in terms of achieving the best treatment multipurpose outcomes. We use this approach in the task of creating a practical tool for finding optimal therapy for T2DM disease. The proposed tool was validated using surveys of experts, clinical recommendations [1], and classic metrics for predictive task. All these validations have shown that the proposed tool is high-quality, interpretable and usability, therefore it can be used as part of the Decision Support System for medical specialists who work with T2DM patients.

Keywords: optimal therapy, predictive modeling, expert-based modeling, hybrid approach, diabetes mellitus, machine learning.

1 Introduction

In modern practical medicine, there are many approaches to personalize recommending therapy to a patient. Medical experts without experience selecting therapy based on clinical guidelines [1] for the treatment of a specific disease. However, clinical guidelines cannot consider the whole variety of combinations of patient conditions and combinations of drugs. More experienced specialists select therapy based on their own experience, studies, and fundamental knowledge of the particular disease course. However, the combinations space of indicators with individual patient's treatment history is multidimensional and multicomponent. The experience of experts may be insufficiently to make a decision in each specific case from this space. Therefore, special methods are required for making decisions in selecting therapy tasks. In this work, we present a new hybrid approach for finding the optimal therapy based on statistical modeling and modeling of the dynamics of the course of the disease.

ICCS Camera Ready Version 2021
To cite this paper please use the final published version:
DOI: 10.1007/978-3-030-77967-2_40

2 Related works and problem definition

Methods for solving the problem of finding optimal therapy for chronic disease are widely described in the literature. These methods include 3 approaches.

Articles in the first approach describe methods for identifying patterns of the effect of a particular therapy on specific treatment targets. Patterns are identified using statistical modeling tools such as statistical hypothesis testing and correlations. In work [2] authors discusses the principles of rational using of antibiotics for sepsis and septic shock and presents scientifically based recommendations for optimal antibiotic therapy. In this work [3], experts studied the effect of two diabetic drugs on blood composition and calculated the coefficients of the effectiveness of these drugs. Jason K. At al. show the advantages of personalizing selection of cancer therapy in the work [4]. Burgmaier et al. have reviewed the potential action drugs on cardiovascular disease and summarize the potential role of present glucagon-like peptide-1-based therapies from a cardiologist's point of view [5]. The methods from this approach are not applicable for personalized selection of the optimal therapy for a particular case. However, using these methods, it is possible to identify patterns that can be a basis for creating methods for finding the optimal therapy.

The second approach includes articles describing methods of identifying linear or non-linear patterns between particular drugs and treatment-goal indicators. Tools for identifying patterns includes Machine Learning methods [6], [7] including Deep Learnings using neural networks [8]. Menden M. and al. predict the response of cancer cell lines to drug treatment. Models predicted IC50 values by 8-fold cross-validation and an independent blind test with coefficients of determination R2 0.72 and 0.64, respectively [6]. In this study [7], Khaledi A. and al. sequenced the genomes and transcriptomes of 414 drug-resistant clinical Pseudomonas aeruginosa isolates. Researchers generated predictive models and identified biomarkers of resistance to four commonly administered antimicrobial drugs. In the work [8] Barbieri C. and al. use feedforward artificial neural network for predicting the response to anemia treatment. Using this approach experts can predict effectiveness of particular drugs. However, a lot of this articles haven't proposed method for finding the most effective therapy in drugs combinations form.

In contrast to the second approach, the third approach includes methods for predicting the synergy of new drug combinations. This approach is based on special methods, such as Tree Combo [9] , Deep Synergy [10], and also Machine Learning methods [11]. Janizek J. and al. introduce new extreme gradient boosted tree-based approach to predict synergy of novel drug combinations, using chemical and physical properties of drugs and gene expression levels of cell lines as features [10]. The second examples of this approach is work [10], that describes method of predicting drugs synergy based on Deep Learning. This method was compared with other machine learning methods, DeepSynergy significantly outperformed the other methods with an improvement of 7.2% over the second-best method at the prediction of novel drug combinations within the space of explored drugs and cell lines. In the work [12] Kuenzi B. and al. developed DrugCell, an interpretable deep learning model of human cancer cells trained on the responses of 1,235 tumor cell lines to 684 drugs. Analysis of the DrugCell results leads to the development of synergistic drug combinations that are validated using combinatorial CRISPR, in vitro drug-drug screening, and patient-derived xenografts. DrugCell

ICCS Camera Ready Version 2021
To cite this paper please use the final published version:
DOI: 10.1007/978-3-030-77967-2_40

provides a blueprint for constructing interpretable predictive medicine models. Using models from this approach we can predict the synergy of drugs combinations. However, the space of drugs combinations and goal-treatment is multidimensional and multicomponent.

Also, there are a lot of novelties in the pharmacological sphere, and it is necessary to consider those in our task. Butler at all. described new knowledge and new developments in the pharmacological sphere for diabetes treatment[13].

Special methods are needed to find the optimal combination of drugs based on the assessment of the effectiveness of a particular combination (which this approach can predict). Therefore, a new method of searching for optimal therapy is needed. It should include all the advantages of the above methods. In this work, we propose a hybrid method that includes methods of the first and second approaches to identify the relationship between a drug and a target indicator of treatment, methods of the third approach to assesses the synergy of drug combinations, and new methods to find the optimal drug combination for a multi-component treatment goal.

3 Hybrid predictive modelling for finding optimal therapy

In the previous items, we explained necessarily to create an approach for personalized finding optimal therapy.

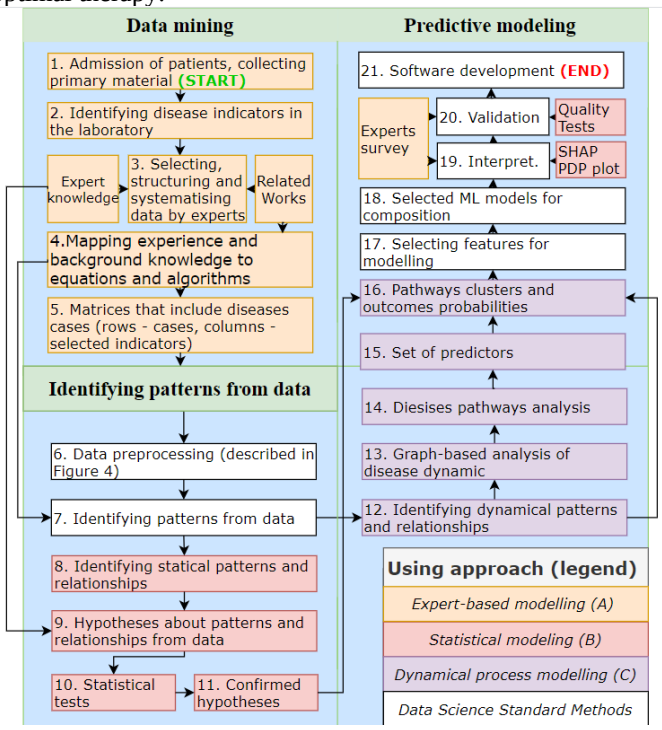

Fig. 1. Hybrid modeling scheme

In this work we propose new hybrid approach for this goal. Scheme of creating this approach is shown in Figure 1.

ICCS Camera Ready Version 2021
To cite this paper please use the final published version:
DOI: 10.1007/978-3-030-77967-2_40

The first stage is data mining. We collaborate with medical experts to transform real-word medical processes into digital form. Primary information includes data from patient's survey, laboratory results, records from electronic medical cards, medical images and other. Medical experts structuring, selecting, and aggregating information for hypothesis and modelling. Our team studying related works and mapping experience and background knowledge to equations, algorithms, and digital patterns. Next, we create special scripts for transforming information by medical experts to matrices. Rows are cases of diseases; columns are important indicators of course of the disease. In this stage of method, we use only expert-based modelling. This stage includes 1th-5th steps.

The second stage is identification of patterns from data. This stage includes steps 6th-10th from scheme on Figure 1 (further just a scheme). In the first, data are preprocessed, it is 6th steps. This step includes the following: noise and emission processing, removing/replacing data gaps, coding categorical features using one-hot-label-encoding/dummy encoding methods, scaling, and logging. Next, in the 8th step, we are identifying statical patterns from data. In parallel, we are using dynamical process modelling for identifying dynamical patterns and relationships in the 12th steps.

The third stage is predictive modelling. This stage includes steps 11th-17th from scheme. We use dynamical and statical patterns for selecting indicators for final features set. All samples are divided into training and testing parts. Next, we create ML predictive models. Models' selection do use cross-validation for only training samples.

Next, we create optimizing treatment goals method. For this we create set of synthetic therapies in form of random combination of drugs. Set include a lot of variants of drugs. Then, we find top-100 best variants and creating only one drugs combination using this combination. This method describes in item 4.3 Optimization**.** Next, we upgrade this method using dynamical patterns, it is step 13th.

Next step is interpretation. Interpretation methods include Shapley Additive explanations (SHAP), Partial Depends Plots (PDP) and expert-based interpretation. Then, we validate methods using expert's surveys, classic metric of predictive task quality and comparing results of using recommended therapy with real results.

To summary, hybrid predictive modelling includes several approaches for finding optimal therapy – expert-based modeling, statistical modeling, and dynamic process modeling. We demonstrate this approach using the case study of finding optimal therapy for treatment diabetes mellitus of two type (T2DM).

4 Finding optimal therapy: T2DM case-study

4.1 Problem description: T2DM-study

Diabetes mellitus (DM) is one of the most common chronic diseases in the world. Experts from the World Diabetes Federation predict, predict that the number of patients with diabetes by 2030 will increase 1.5 times and reach 552 million people, mainly due to patients with type 2 diabetes (T2DM). For public health, this type of diabetes is one of the most priority problems, since this disease is associated with a large number of concomitant diseases, leading to early disability, and increased cardiovascular risk. Therefore, it is especially important for patients with this disease to prevent the

ICCS Camera Ready Version 2021
To cite this paper please use the final published version:
DOI: 10.1007/978-3-030-77967-2_40

development of serious complications. The risks of complications can be reduced by right selected therapy, but it is important to choose the right drugs, considering the synergy of drugs and the patient's personal characteristics. There are many brands to treat this disease. Treatment targets are multicomponent and include carbohydrates metabolism compensation (glycated hemoglobin), lipid control (total cholesterol), and optimization of systolic and diastolic pressure-level. The space of drug combinations and treatment targets is multidimensional and multi-component, special methods are required for personalized searching drug combinations, that are optimal in terms of better treatment multi-component outcomes. Therefore, the task of creating a method for a personalized search for optimal therapy for patients with type 2 diabetes is relevant and suitable for demonstrating the proposed hybrid approach.

4.2 Data mining

The study was based on dataset including 189 671 medical records for patients who were treated for diabetes type 2 in Almazov National Medical Research Centre or in Pavlov First Saint Petersburg State Medical University, St. Petersburg, Russia in 2008-2018. There are several entry and exclusionary criteria's for including treatment case in study. Criteria showed in Table 1.

Table 1. Entry and exclusionary criteria for including treatment case in dataset.

Entry criteria	Exclusionary criteria
1. Diabetes type 2 2. At least 2 measurements of glycated hemoglobin 3. Age between 18 and 80 years	1. Early stages of diabetes, prediabetes, impaired glucose tolerance 2. The presence of a large number of gaps in key-indicators 3. The observation period is less than half a year

Each treatment case describes using set of indicators. This is shown in Table 2.

Table 2. Medical indicators for treatment cases

Feature's group	Features
Measurements	Group includes height, weight, age, gender, SBP, DBP, pulse, body mass index, body surface area;
Hypertensions	Group includes i10, i11, i12, i13, i15 ICD codes;
Heart complications	Group includes chronic heart failure, chronic obstructive pulmonary disease, atherosclerosis, myocardial infarction, acute coronary syndrome, and others;
Diabetic complications	Group includes retinopathy, angiopathy, nephropathy, neuropathy, foot ulcer, diabetic coma, osteoarthropathy and others;
Other nosologies	Group includes anemia, hypothyroidism, acute pulmonary complications (pneumonia, bronchitis, other types of pneumonia);
Insulin	Group includes 117 insulin preparations, short, medium, long-acting insulins, genetically engineered insulins, and many others:

ICCS Camera Ready Version 2021
To cite this paper please use the final published version:
DOI: 10.1007/978-3-030-77967-2_40

Sugar-lowering drugs	Group includes 213 types of drugs. These drugs are most often used in the treatment of diabetes in the Russian Federation. This drug includes metformin, diabetalong, lixumia, and others;
Other drugs	Group includes 211 different types of diuretics (e.g., Aldacton Saltucin), 108 different types of statins (e.g., Anvistat), 193 types of beta-blockers (e.g., Normoglaucon).

These medical indicators include information about diseases in anamnesis, analysis' values, physical measurements, lifestyle, a lot of drug types.

4.3 Optimization

Data mining. The data were presented as a time series of treatment of all patients. However, this data was not suitable for processing by traditional methods like ARIMA or LSTM network. It was caused by availability of a large number of data gaps. The second reason for not using these methods was the presence of various distances between visits. Due to these reasons series for all patients were compressed into one-dimensional vector, which represents series with its statistical characteristics. The compression scheme showed in Figure 2.

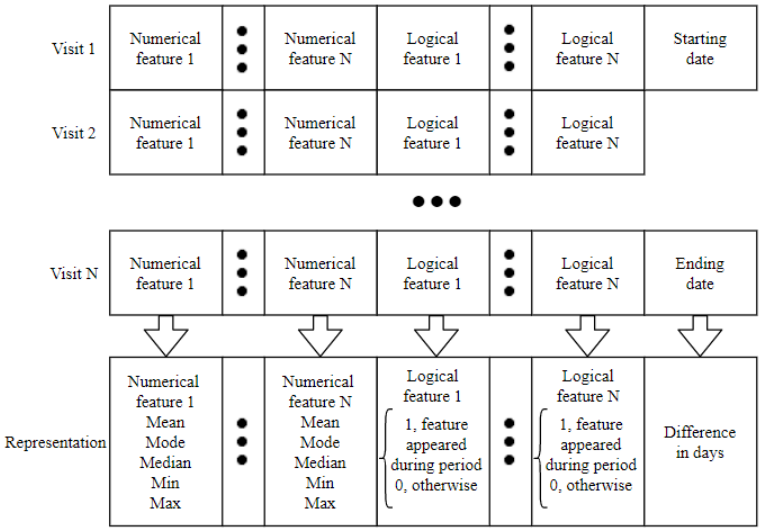

Fig. 2. Time series transformation scheme

This transformation was applied for visits which were between any two measurements between target values. As targets for predicting were chosen 4 features, they are glycated hemoglobin, total cholesterol, systolic and diastolic pressure. For targets were applied other transformations, for glycated hemoglobin and total cholesterol the difference between the end and the start values in the series was calculated. However, for systolic and diastolic pressure this operation should not be applied, since these features change throughout the day, then the use of one value will be incorrect. Therefore, for pressures were calculated mean values within six months after the treatment.

ICCS Camera Ready Version 2021
To cite this paper please use the final published version:
DOI: 10.1007/978-3-030-77967-2_40

Predictive modeling. Since the main goal is to reduce key indicators, first of all it is necessary to be able to predict their values by using treatment as a predictor. Thus, the problem becomes building a regression model.

Machine learning models were created for all key indicators. The following algorithms were used: Decision Tree, Random Forest, XGBoost, SGD, CatBoost. Mean squared error and coefficient of determination were used as validation metrics. Metrics of trained models showed in the Tables 3-6.

Table 3. Metrics of glycated hemoglobin predicting model.

Model	**Mean squared error (CI=95%)**	**R^2 (CI=95%)**
Decision Tree	(0.2220-0.5290) 0.3755	(0.4270-0.7616) 0.5943
Random Forest	(0.1518-0.3134) 0.2326	(0.6792-0.8338) 0.7565
XGBoost	(0.1312-0.2784) 0.2048	(0.7042-0.8594) 0.7818
SGD	(0.5104-0.7710) 0.6407	(0.2412-0.4118) 0.3265
CatBoost	(0.1330-0.2546) 0.1938	(0.7312-0.8594) 0.7953

Table 4. Metrics of total cholesterol predicting model.

Model	**Mean squared error (CI=95%)**	**R^2 (CI=95%)**
Decision Tree	(0.6806-0.9546) 0.8176	(0.3278-0.5294) 0.4286
Random Forest	(0.4086-0.5642) 0.4864	(0.6160-0.7076) 0.6618
XGBoost	(0.4106-0.5690) 0.4898	(0.6080-0.7094) 0.6587
SGD	(1.0048-1.3264) 1.1656	(0.1420-0.2414) 0.1917
CatBoost	(0.4214-0.5712) 0.4963	(0.6092-0.6986) 0.6539

Table 5. Metrics of systolic pressure predicting model.

Model	**Mean squared error (CI=95%)**	**R^2 (CI=95%)**
Decision Tree	(24.4212-29.8376) 27.1294	(0.9072-0.9236) 0.9154
Random Forest	(15.4982-18.5108) 17.0045	(0.9424-0.9516) 0.9470
XGBoost	(31.6220-35.1730) 33.3975	(0.8908-0.9010) 0.8959
SGD	(190.2478-203.8474) 197.0476	(0.3756-0.3962) 0.3859
CatBoost	(32.2876-35.7460) 34.0168	(0.8892-0.8990) 0.8941

Table 6. Metrics of diastolic pressure predicting model.

Model	**Mean squared error (CI=95%)**	**R^2 (CI=95%)**
Decision Tree	(17.8134-21.7346) 19.7740	(0.9068-0.9232) 0.9150
Random Forest	(10.9792-13.0486) 12.0139	(0.9438-0.9528) 0.9483
XGBoost	(23.7260-26.0768) 24.9014	(0.8886-0.8972) 0.8929
SGD	(161.9712-171.8524) 166.9118	(0.2708-0.2918) 0.2813

ICCS Camera Ready Version 2021
To cite this paper please use the final published version:
DOI: 10.1007/978-3-030-77967-2_40

CatBoost	(24.5408-26.8876) 25.7142	(0.8850-0.8938) 0.8894

Since trained models could predict our target values it is possible to change medicine, which was used in treatment and to see how key indicator changed. For choosing best medicine combination it is necessary to apply models for all drugs combinations. However, there are 87 different medicaments in the dataset, thus 2^{87} combinations should be checked, which is time-consuming task.

Treatment selection becomes an optimization problem, where our target to find vector of medicines, which gives most acceptable values of key indicators. There are several difficulties with solving this problem. First problem is that our target function is a black box function, which means that we do not know its behavior. The second one is that it has high evaluation cost, consequently using metaheuristic optimization algorithms will be time consuming. The third problem is that it is necessary to pick up only integer values, because vector of medicines consists of zeros and ones.

Since no suitable method was found to solve the optimization problem, the following approach was suggested. Scheme of this approach showed in Figure 4.

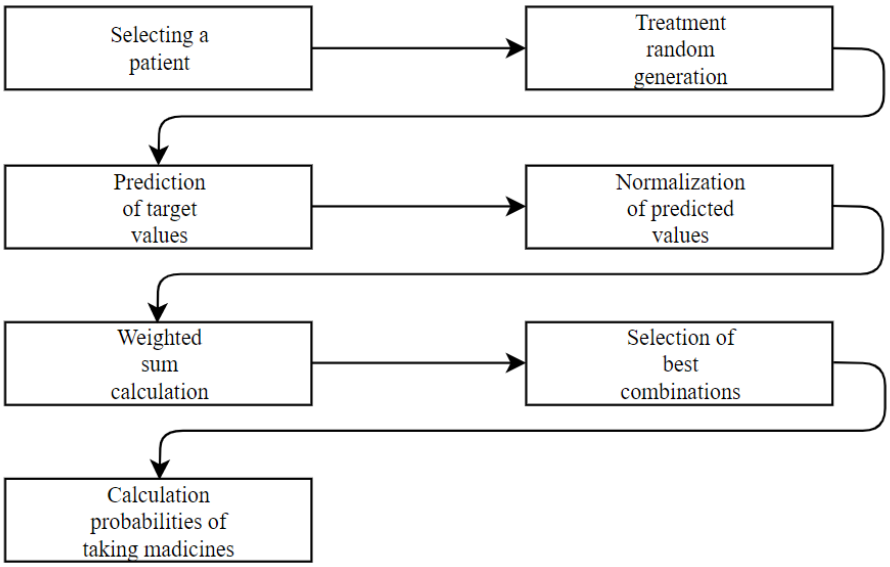

Fig. 4. Drug selection scheme

For a patient randomly generates 100000 different combinations of drugs. Next step is to predict key indicators values with ML models. Then the difference between the predicted value and the value that the patient needs to obtain to get within the acceptable interval is calculated. Next it is necessary to normalize these differences, multiply by the coefficient of importance and summarize them. Then we need to select the top 100 obtained values with their corresponding medicines combinations and use them to calculate the probability of taking the drug into treatment. For each drug we calculate its frequency of occurrence in top 100 combinations and divide it by 100. The last step is

ICCS Camera Ready Version 2021
To cite this paper please use the final published version:
DOI: 10.1007/978-3-030-77967-2_40

the selection of drugs according to a random number generator based on the obtained probabilities.

Interpreting. With using SHAP values it is possible to interpret model output, which means that an importance of concrete drug and its influence could be found out. SHAP values of glycated hemoglobin prediction model showed in Figure 5.

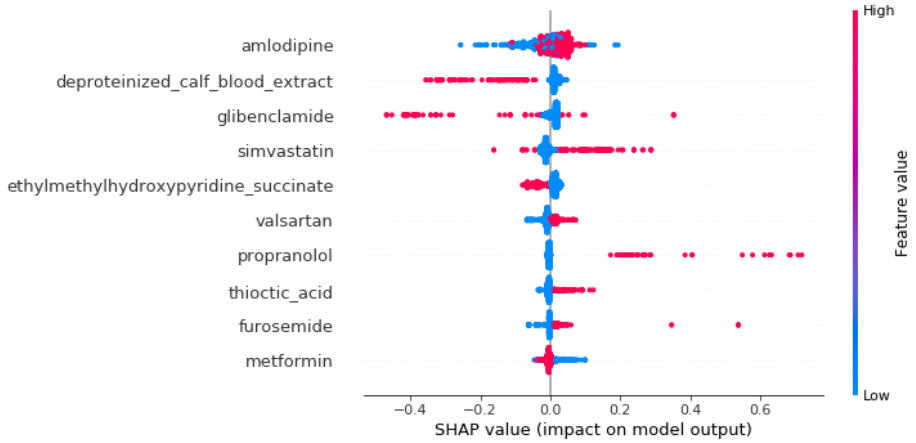

Fig. 5. SHAP values of glycated hemoglobin prediction model

The graph shows the 10 drugs most affecting glycated hemoglobin. Red color means that the medicine was included in the treatment, blue is the opposite. According to a model deprotenized calf blood extract and glibenclamide are most effective drugs in terms of the lowering the indicator. On the other hand, simvastatin and propranolol raise this value most effectively. Also, there is interesting case with metformin, it does not decrease target value, but without using it this value is increasing, which mean that it could be used for keeping the indicator in appropriate range.

SHAP values of systolic pressure prediction model showed in Figure 6.

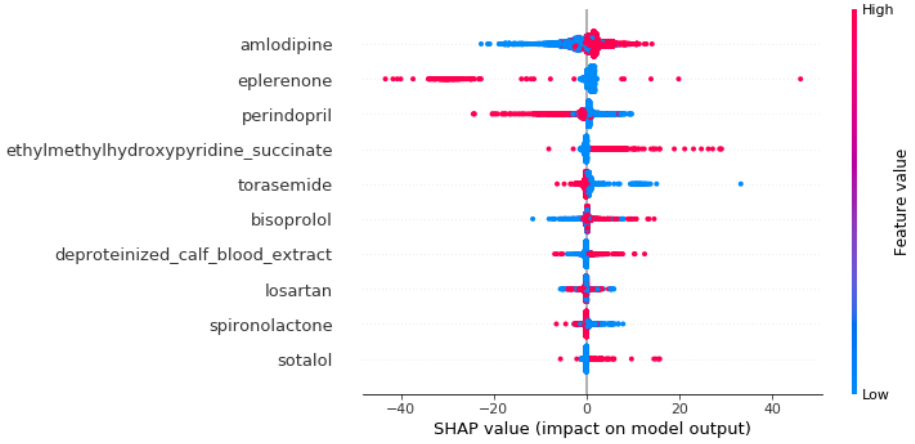

Fig. 6. SHAP values of systolic pressure prediction model

According to this plot ethylmethyldroxypytidine succinate is most systolic pressure affective in terms of increasing this value. However, this medicine also in the list of

ICCS Camera Ready Version 2021
To cite this paper please use the final published version:
DOI: 10.1007/978-3-030-77967-2_40

most glycated hemoglobin affective drugs, but it lowers this value. This is a good example of why it is important to choose the right treatment, because bad combination could optimize only one target to the detriment of another. Therefore, patients taking this drug should also be prescribed medication to compensate for the increase in systolic pressure.

SHAP values of diastolic pressure prediction model showed in Figure 7.

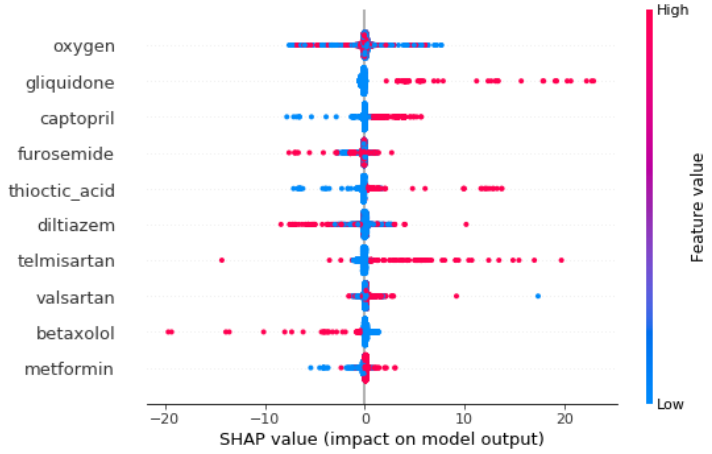

Fig. 7. SHAP values of diastolic pressure prediction model

Graph above shows good example of drug, which can be used to maintain acceptable values of key indicators. This is metformin, when using this drug, there is no need to compensate for the change in any indicator.

SHAP values of total cholesterol prediction model showed in Figure 8.

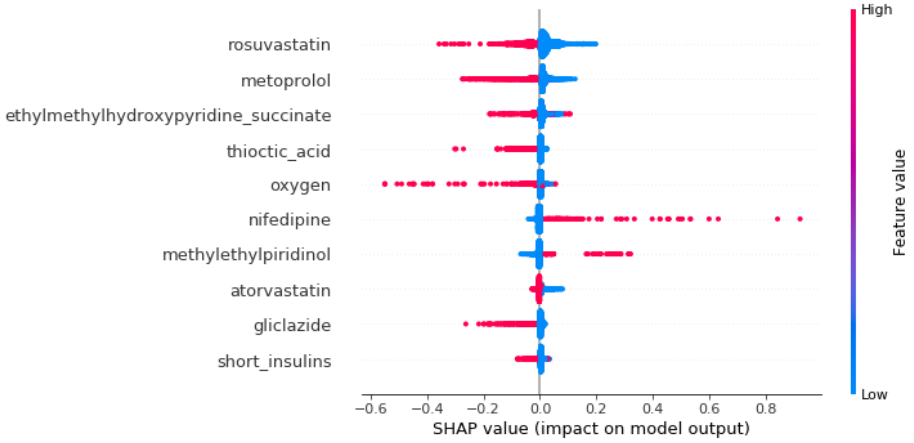

Fig. 8. SHAP values of total cholesterol prediction model

From this graph we can see that using medical oxygen greatly decreases total cholesterol value, but there is a problem with using this drug. SHAP values of this drug from diastolic pressure table are hard to interpret. It means that if this medication will

ICCS Camera Ready Version 2021
To cite this paper please use the final published version:
DOI: 10.1007/978-3-030-77967-2_40

be included into a treatment, it will be hard to compensate increasing or decreasing of diastolic pressure, since changes in target value can be in any direction.

Validating. The optimization scheme was applied to patients for whom it was possible to calculate all 4 key indicators. Results showed in table 7.

Table 7. Optimization scheme validation results.

Target value	Percentage of cases, when target value brought back to normal range	Percentage of cases, when target value became better that in real treatment
Glycated hemoglobin	67.64	**91.17**
Total cholesterol	20.58	**26.47**
Systolic pressure	47.05	**82.35**
Diastolic pressure	64.70	**79.41**
GH + TC	20.58	26.47
GH + SP	26.47	73.52
GH + DP	50.00	70.58
TC + SP	17.64	20.58
TC + DP	14.70	17.64
SP + DP	26.47	70.58
GH + TC + SP	17.64	20.58
GH + TC + DP	14.70	17.64
GH + SP + DP	20.58	61.74
TC + SP + DP	11.76	17.64
GH + TC + SP + DP	11.76	17.64

According to given results, most important target, which is glycated hemoglobin, becomes much better than in real cases, systolic and diastolic pressure also becomes better in big number of cases, however total cholesterol value is poorly improved. To increase the number of cases of improvement in total cholesterol, it is necessary to improve the model.

ICCS Camera Ready Version 2021
To cite this paper please use the final published version:
DOI: 10.1007/978-3-030-77967-2_40

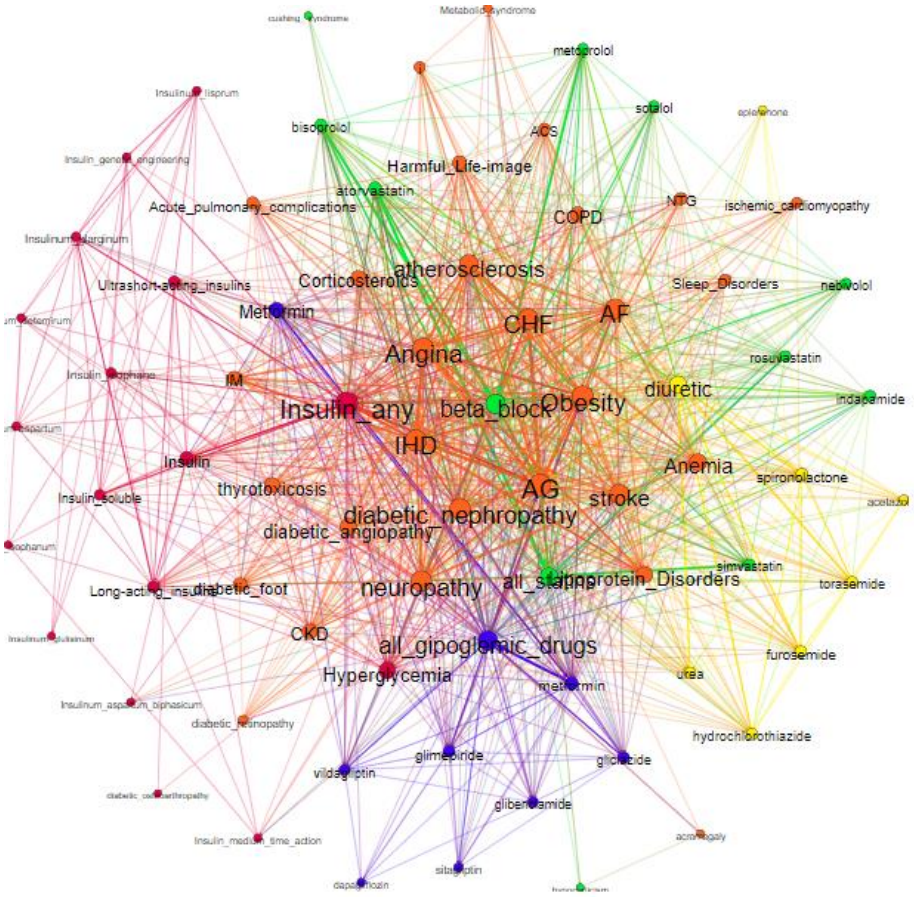

Fig. 9. Graph of therapy trajectories. Colors are marks for clusters by modularity maximization.

We improve this model using dynamic patterns of the treatment process. For this, we analyze space of therapy trajectories (Figure 9 shows this space). For each patient, its treatment trajectory is identified in the form of a sequence of diagnosed certain complications from the past history of the disease. Then, using the Tanimoto coefficient [14], we find similar trajectories (patients with similar medical past histories). Further, for these patients, we determine a set of pharmacological groups of drugs that show the best results for optimizing a multipurpose result. To determine this set of pharmacological groups, we use our developed algorithm based on the analysis of precedents. Further, we apply the above method, choosing drugs not from 87 drugs, but from a reduced number of drugs from individually selected pharmacological groups of drugs. Table 8 shows the change in the quality indicators of the model from Table 7.

ICCS Camera Ready Version 2021
To cite this paper please use the final published version:
DOI: 10.1007/978-3-030-77967-2_40

Table 8. Result of improve model for increasing number improving total cholesterol.

Target value	Percentage of cases when target value brought back to normal range	Percentage of cases when target value became better that in real treatment
Total cholesterol	-	44.11 **(+17,64%)**
GH + TC	94.11 **(+73,53%)**	41.17 **(+14,7%)**
TC + SP	26.47 **(+9%)**	29.41 **(+9%)**
TC + DP	-	20.58 **(+2,94%)**

5 Conclusion and Future Work

This paper proposes a hybrid approach for creating a method for finding optimal therapy in the terms of optimizing the multipurpose outcome of patient treatment. This approach is based on identifying statistical and dynamic patterns from a course of the disease using expert-based modelling methods, machine learning methods, and predictive modelling methods. This method was demonstrated on the practical task of finding the optimal therapy for type 2 diabetes mellitus in terms of achieving 4 treatment goals - compensation of carbohydrate metabolism (target value is glycated hemoglobin), compensation of lipid metabolism (target value is total cholesterol), optimization of half-year indicators of arterial pressure (target value are systolic and diastolic). This method was validated using a survey of experts-endocrinologists, classical metrics of predictive modelling tasks. Also, we have validated this method using real-treatments cases. We have found the optimal therapy for each case and have predicted results of the optimal therapy for each case using developed predictive models. As a result, the proposed hybrid method improves the target indicators of carbohydrate metabolism compensation in 91% of cases, the target of lipid metabolism in 44% of cases, the average semi-annual systolic pressure in 82% of cases, and average semi-annual diastolic pressure in 80% of cases, compared with real therapy for the selected patients. In summary, the method is of high quality, it can be applicated as part of a support and decision-making system for medical specialists working with T2DM patients.

Acknowledgments. The reported study was funded by RFBR according to the research project № 20-31-70001. Participation in the ICCS conference was supported by the NWO Science Diplomacy Fund project # 483.20.038 "Russian-Dutch Collaboration in Computational Science"

References

[1] I. I. Dedov *et al.*, "Standards of specialized diabetes care. Edited by Dedov I.I., Shestakova M.V., Mayorov A.Yu. 9th edition," *Diabetes Mellit.*, 2019.

[2] A. Kumar, "Optimizing Antimicrobial Therapy in Sepsis and Septic Shock," *Critical Care Clinics*. 2009.

[3] P. Srividya, T. S. R. Devi, and S. Gunasekaran, "Ftir Spectral Study on Diabetic Blood Samples – Monotherapy and Combination Therapy," *Ojp*, 2012.

ICCS Camera Ready Version 2021
To cite this paper please use the final published version:
DOI: 10.1007/978-3-030-77967-2_40

[4] J. K. Sicklick *et al.*, "Molecular profiling of cancer patients enables personalized combination therapy: the I-PREDICT study," *Nat. Med.*, 2019.

[5] M. Burgmaier, C. Heinrich, and N. Marx, "Cardiovascular effects of GLP-1 and GLP-1-based therapies: Implications for the cardiovascular continuum in diabetes?," *Diabetic Medicine*. 2013.

[6] M. P. Menden *et al.*, "Machine Learning Prediction of Cancer Cell Sensitivity to Drugs Based on Genomic and Chemical Properties," *PLoS One*, 2013.

[7] A. Khaledi *et al.*, " Predicting antimicrobial resistance in Pseudomonas aeruginosa with machine learning-enabled molecular diagnostics ," *EMBO Mol. Med.*, 2020.

[8] C. Barbieri *et al.*, "A new machine learning approach for predicting the response to anemia treatment in a large cohort of End Stage Renal Disease patients undergoing dialysis," *Comput. Biol. Med.*, 2015.

[9] J. D. Janizek, S. Celik, and S. I. Lee, "Explainable machine learning prediction of synergistic drug combinations for precision cancer medicine," *bioRxiv*. 2018.

[10] K. Preuer, R. P. I. Lewis, S. Hochreiter, A. Bender, K. C. Bulusu, and G. Klambauer, "DeepSynergy: Predicting anti-cancer drug synergy with Deep Learning," *Bioinformatics*, 2018.

[11] A. A. Tabl, A. Alkhateeb, H. Q. Pham, L. Rueda, W. ElMaraghy, and A. Ngom, "A Novel Approach for Identifying Relevant Genes for Breast Cancer Survivability on Specific Therapies," *Evol. Bioinforma.*, 2018.

[12] B. M. Kuenzi *et al.*, "Predicting Drug Response and Synergy Using a Deep Learning Model of Human Cancer Cells," *Cancer Cell*, 2020.

[13] J. Butler, J. L. Januzzi, and J. Rosenstock, "Management of heart failure and type 2 diabetes mellitus: Maximizing complementary drug therapy," *Diabetes, Obesity and Metabolism*. 2020.

[14] D. Bajusz, A. Rácz, and K. Héberger, "Why is Tanimoto index an appropriate choice for fingerprint-based similarity calculations?," *J. Cheminform.*, 2015.

ICCS Camera Ready Version 2021
To cite this paper please use the final published version:
DOI: 10.1007/978-3-030-77967-2_40