Preface
The 5th International Workshop on Knowledge Discovery in Healthcare Data (KDH)

Introduction

The Knowledge Discovery in Healthcare Data (KDH) workshop series was established in 2016 to bring together AI and clinical researchers, fostering collaborative discussions and presenting AI research efforts to solve pressing problems in health care. This fifth edition of the workshop was held in conjunction with the 24th European Conference on Artificial Intelligence, Digital ECAI 2020, which was hosted in Santiago de Compostela, Spain, but conducted virtually. The focus of the workshop was on learning health care systems. For the second time, this workshop featured a challenge: The Blood Glucose Level Prediction (BGLP) Challenge.

The notion of the learning health care system has been put forward to denote the translation of routinely collected data into knowledge that drives the continual improvement of medical care. This notion has been described in many forms, but each follows a similar cycle of assembling, analyzing and interpreting data from multiple sources (clinical records, guidelines, patient-provided data including wearables, omic data, etc.), followed by feeding the acquired knowledge back into clinical practice. This framework aims to provide personalized recommendations and decision support tools to aid both patients and care providers, to improve outcomes and personalize care.

This framework also extends the range of actions possible in response to patient monitoring data, for example, alerting patients or automatically adjusting insulin doses when blood glucose levels are predicted to go out of range. Blood glucose level prediction is a challenging task for AI researchers with the potential to improve the health and well-being of people with diabetes. In the Blood Glucose Level Prediction (BGLP) Challenge, researchers came together to compare the efficacy of different machine learning (ML) prediction approaches on a standard set of real patient data.

The workshop received 35 submissions, each of which was peer-reviewed by three reviewers. Based on the reviews, 10 technical papers and 16 BGLP Challenge papers were accepted for presentation at the workshop. Among the accepted papers, the current trend of applying deep learning (DL) is strongly represented, while other methods used are case-based reasoning (CBR), natural language processing, and time series analysis. Another evident trend was the need for open data sets that can drive the field forward and promote building on each other's work. This topic was addressed by the invited talk as well as by the included BGLP Challenge.

Keynote Speaker: Kerstin Bach, NTNU, Norway

Bio: Kerstin Bach is an Associate Professor of Computer Science and Artificial Intelligence in the Department of Computer Science at the Norwegian University of Science and Technology (NTNU). She has been at NTNU since 2017, where she is currently deputy head of the Data and Artificial Intelligence group and a core member of the Norwegian Open AI Lab. Bach received her doctorate summa cum laude from the Department of Mathematics, Natural Sciences, Economics and Computer Science of the Hildesheim University, Germany, in 2012.

Kerstin Bach has broad experience building industrial strength AI applications as well as leading and collaborating on interdisciplinary teams. While working at Verdande Technology, she worked on a platform delivering AI services for the Oil and Gas, Finance and Healthcare sector. Further, she has headed the myCBR open source project since 2010 and has conducted research projects leveraging CBR and other AI methods for over 13 years. She is currently focused on two Horizon 2020 projects, selfBACK and AI4EU. She is the project manager of the selfBACK project, responsible for the technical integration of selfBACK into Back-UP, where she leads the Machine Learning tasks. In the AI4IoT pilot of AI4EU, she co-leads the efforts to develop AI showcases for the platform featuring Air Quality measurements. Bach is active in communicating AI research internationally. She is the chair of the German Special Interest Group on Knowledge Management and a board member of the Norwegian AI Society.

Title: The Potential for AI in Public Health: Lessons Learned from Developing and Testing a Patient-Centered Mobile App

Abstract: This talk provides an overview of how Artificial Intelligence and Machine Learning have been used to develop a mobile app that facilitates self-management of low back pain patients. It covers the development of the decision support system for patients using case-based reasoning as well as system evaluation via a randomized controlled trial testing the effectiveness of the app. This talk focuses on the development of the selfBACK system [24], but the approaches and methodologies employed can also be applied to the development of systems for other chronic diseases benefiting from self-management.

Accepted Papers

Main Track Papers

Main track technical papers present original research work across a broad range of KDH topics and domains. Given the current Covid-19 pandemic, this proceedings features three papers addressing the use of AI for detecting anomalies in X-ray scans. Paper [16] presents an approach for quantifying the uncertainty of deep neural networks (DNN) for the task of chest X-ray image classification, with results showing that utilizing uncertainty information may improve DNN performance for some metrics and observations. Paper [10] presents a study and a concrete tool based on machine learning to predict the prognosis of hospitalized patients with Covid-19. Paper [12] proposes a two-stage segmentation method which is capable of improving the accuracy of detection and segmentation of lung nodules from 2D CT images, achieving promising results that put the method among the top lung nodule segmentation methods.

The second group of papers focuses on how AI-based explanation and visualization can help patients and clinicians use the vast amount of information available to improve diagnosis, knowledge discovery and care. Paper [25] presents InterVENE, an approach that visualizes neural embeddings and interactively explains this visualization, aiming for knowledge extraction and network interpretation. Paper [7] makes use of the graphical representation capabilities of Formal Concept Analysis (FCA) and use graph databases as a visualization method for knowledge patterns. The authors exemplify their approach on a particular medical dataset, highlighting a 3D representation of conceptual hierarchies by using virtual reality. Paper [4] is a position paper, in which the authors analyze the cause-effect relationships for determining the causal status among a set of events. They argue that causal knowledge graphs can improve the accuracy and reliability of existing ML/DL-based diagnosis methods, by producing transparent justifications and explanations of the output. Paper [23] presents initial findings towards assessing how computer vision, natural language processing and other systems could be correctly embedded in the clinicians' pathway to better aid in fracture detection.

A third group of papers addresses the use of machine learning for blood glucose level prediction (BGLP) and diabetes management. Paper [22] compares the effectiveness of several BGLP models and found that Lasso regression performed best out of the algorithms used for both the 30-minute and 60-minute prediction horizons. Paper [1] presents a generic neural architecture previously used for BGLP in a what-if scenario that can be adapted and leveraged to make either carbohydrate or bolus recommendations. Paper [17] addresses the problem of missing sensor readings in glucose monitoring data of artificial pancreas (AP) systems. It uses data from virtual patients and a state-of-the-art AP controller simulating various scenarios.

BGLP Challenge Papers

The BGLP Challenge papers describe blood glucose (BG) level prediction approaches and experimental evaluations on the newly updated OhioT1DM dataset [20]. Of the 16 systems with papers that were accepted for publication, 8 systems had results that conformed to The BGLP Challenge Rules[1]. These 8 systems were all evaluated using the exact same test points for each of 6 data contributors in the OhioT1DM dataset. Results were reported as the root mean squared error (RMSE) and the mean absolute error (MAE) scores for the 30 minute and 60 minute prediction horizons. The 4 scores were added together to compute an overall score, and the 8 systems were ranked in increasing order of this total score. Table 1 shows the official ranking of the 8 systems, based on this overall score. Additional rankings, e.g. based on each of the 4 measures separately, as well as links to the source code for all 16 systems, are available on The BGLP Results [2] page.

	30 minutes		60 minutes		
Paper	RMSE	MAE	RMSE	MAE	Overall
[29]	18.22	12.83	31.66	23.60	86.31
[11]	19.21	13.08	31.77	23.09	87.15
[32]	18.34	13.37	32.21	24.20	88.12
[31]	19.05	13.50	32.03	23.83	88.41
[2]	18.23	14.37	31.10	25.75	89.45
[30]	19.37	13.76	32.59	24.64	90.36
[14]	19.60	14.25	34.12	25.99	93.96
[19]	20.03	14.52	34.89	26.41	95.85

Table 1: BGLP Challenge overall ranking.

Gated versions (LSTMs [13], GRUs [6]) of recurrent neural networks (RNNs) were predominant, used either at the core of the forecasting model [2, 3, 5, 11, 21], or as a component in a larger model [26, 29]. Other types of neural architectures that were frequently used were convolutional RNNs (CRNNs) [3, 8, 9] and fully connected networks (FCNs) [2, 26, 28]. Generative Adversarial Networks (GANs) were used in [32], wherein the GRU-based generator uses real data as input and its BG predictions are pitted against the true BG values in a discriminator implemented using one-dimensional convolutional neural networks (CNNs). The recently proposed Neural Basis Expansion for Interpretable Time-Series Forecasting (N-BEATS) architecture [27] served as the basis for the winning entry [29]. In this top-performing model, the fully connected block structure of N-BEATS was replaced with LSTMs, additional losses were used to provide more supervision, and secondary, sparse variables such as meals and bolus insulin were used as input while still backcasting only on the primary forecasting variable, blood glucose. A number of non-neural approaches were proposed as well, such as Genetic Programming (GP) for symbolic regression in [14], Random Forests in [14, 28], multivariate Latent Variable (LV) based models in [30], and Partial Least Squares Regression (PLSR) with stacking in [15, 26].

The LSTM-based approach from [5] was notable for its interpretability analysis, wherein the SHAP (SHapley Additive exPlanations) method [18] was used to assess the impact that each feature has on the model predictions. Also of special interest were the "what-if" evaluations from [14], where future values of basal and bolus insulin were assumed to be controlled within the prediction horizon and leveraged with good results in some of the proposed GP-based models. Overall, the participating systems were trained or fine-tuned for each patient (personalized), with the exception of [2] where a single LSTM model was trained to make predictions for all patients (non-personalized).

We very much appreciate the support of the Digital ECAI 2020 workshop chairs, Magdalena Ortiz and Amparo Alonso, as well as this year's general chair Jérôme Lang. Further, we would like to thank Jernej Masnec, of Underline.io, the digital platform provider, for technical support.

We sincerely hope that the participants enjoyed this year's workshop program and that this collection of papers will inspire and encourage more AI-related research for and within healthcare in the future.

Kerstin Bach, Razvan Bunescu,
Cindy Marling and Nirmalie Wiratunga

Santiago de Compostela, virtually, August 2020

[1] http://smarthealth.cs.ohio.edu/bglp/bglp-rules.html
[2] http://smarthealth.cs.ohio.edu/bglp/bglp-results.html

REFERENCES

[1] Jeremy Beauchamp, Razvan Bunescu, and Cindy Marling, 'A general neural architecture for carbohydrate and bolus recommendations in type 1 diabetes management', in *this volume*, (August 2020).

[2] Robert Bevan and Frans Coenen, 'Experiments in non-personalized future blood glucose level prediction', in *this volume*, (August 2020).

[3] Ananth Reddy Bhimireddy, Priyanshu Sinha, Bolu Oluwalade, Judy Wawira Gichoya, and Saptarshi Purkayastha, 'Blood glucose level prediction as time-series modeling using sequence-to-sequence neural networks', in *this volume*, (August 2020).

[4] Eva Blomqvist, Marjan Alirezaie, and Marina Santini, 'Towards causal knowledge graphs - position paper', in *this volume*, (August 2020).

[5] Giacomo Cappon, Lorenzo Meneghetti, Francesco Prendin, Jacopo Pavan, Giovanni Sparacino, Simone Del Favero, and Andrea Facchinetti, 'A personalized and interpretable deep learning based approach to predict blood glucose concentration in type 1 diabetes', in *this volume*, (August 2020).

[6] Kyunghyun Cho, Bart Van Merriënboer, Caglar Gulcehre, Dzmitry Bahdanau, Fethi Bougares, Holger Schwenk, and Yoshua Bengio, 'Learning phrase representations using RNN encoder-decoder for statistical machine translation', in *Proceedings of the 2014 conference on empirical methods in natural language processing (EMNLP)*, pp. 1724–1734, (2014).

[7] Diana Cristea, Christian Sacarea, and Diana Şotropa, 'Knowledge discovery and visualization in healthcare datasets using formal concept analysis and graph databases', in *this volume*, (August 2020).

[8] John Daniels, Pau Herrero, and Pantelis Georgiou, 'Personalised glucose prediction via deep multitask networks', in *this volume*, (August 2020).

[9] Jonas Freiburghaus, Aïcha Rizzotti-Kaddouri, and Fabrizio Albertetti, 'A deep learning approach for blood glucose prediction and monitoring of type 1 diabetes patients', in *this volume*, (August 2020).

[10] Alfonso Emilio Gerevini, Roberto Maroldi, Matteo Olivato, Luca Putelli, and Ivan Serina, 'Prognosis prediction in Covid-19 patients from lab tests and X-ray data through randomized decision trees', in *this volume*, (August 2020).

[11] Hadia Hameed and Samantha Kleinberg, 'Investigating potentials and pitfalls of knowledge distillation across datasets for blood glucose forecasting', in *this volume*, (August 2020).

[12] Mohammad Hesam Hesamian, Wenjing Jia, Sean He, and Paul Kennedy, 'Region proposal network for lung nodule detection and segmentation', in *this volume*, (August 2020).

[13] Sepp Hochreiter and Jürgen Schmidhuber, 'Long Short-Term Memory', *Neural computation*, **9**(8), 1735–1780, (1997).

[14] David Joedicke, Oscar Garnica, Gabriel Kronberger, José Manuel Colmenar, Stephan Winkler, Jose Manuel Velasco, Sergio Contador, and Ignacio Hidalgo, 'Analysis of the performance of genetic programming on the blood glucose level prediction challenge 2020', in *this volume*, (August 2020).

[15] Heydar Khadem, Hoda Nemat, Jackie Elliott, and Mohammed Benaissa, 'Multi-lag stacking for blood glucose level prediction', in *this volume*, (August 2020).

[16] Yumin Liu, Claire Zhao, and Jonathan Rubin, 'Uncertainty quantification in chest X-ray image classification using Bayesian deep neural networks', in *this volume*, (August 2020).

[17] Yunjie (Lisa) Lu, Abigail Koay, and Michael Mayo, 'In silico comparison of continuous glucose monitor failure mode strategies for an artificial pancreas', in *this volume*, (August 2020).

[18] Scott M. Lundberg and Su-In Lee, 'A unified approach to interpreting model predictions', in *Advances in Neural Information Processing Systems 30*, eds., I. Guyon, U. V. Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan, and R. Garnett, 4765–4774, Curran Associates, Inc., (2017).

[19] Ning Ma, Yuhang Zhao, Shuang Wen, Tao Yang, Ruikun Wu, Rui Tao, Xia Yu, and Hongru Li, 'Online blood glucose prediction using autoregressive moving average model with residual compensation network', in *this volume*, (August 2020).

[20] Cindy Marling and Razvan Bunescu, 'The OhioT1DM dataset for blood glucose level prediction: Update 2020', in *this volume*, (August 2020).

[21] Michael Mayo and Tomas Koutny, 'Neural multi-class classification approach to blood glucose level forecasting with prediction uncertainty visualisation', in *this volume*, (August 2020).

[22] Richard McShinsky and Brandon Marshall, 'Comparison of forecasting algorithms for type 1 diabetic glucose prediction on 30 and 60-minute prediction horizons', in *this volume*, (August 2020).

[23] Carlos Francisco Moreno-Garca, Truong Dang, Kyle Martin, Manish Patel, Andrew Thompson, Lesley Leishman, and Nirmalie Wiratunga, 'Assessing the clinicians' pathway to embed artificial intelligence for assisted diagnostics of fracture detection', in *this volume*, (August 2020).

[24] Paul Jarle Mork and Kerstin Bach, 'A decision support system to enhance self-management of low back pain: Protocol for the selfBACK project', *JMIR Res Protoc*, **7**(7), e167, (Jul 2018).

[25] Meike Nauta, Michel van Putten, Marleen C. Tjepkema-Cloostermans, Jeroen Bos, Maurice van Keulen, and Christin Seifert, 'Interactive explanations of internal representations of neural network layers: An exploratory study on outcome prediction of comatose patients', in *this volume*, (August 2020).

[26] Hoda Nemat, Heydar Khadem, Jackie Elliott, and Mohammed Benaissa, 'Data fusion of activity and CGM for predicting blood glucose levels', in *this volume*, (August 2020).

[27] B.N. Oreshkin, D. Carpov, N. Chapados, and Y. Bengio, 'N-BEATS: Neural basis expansion analysis for interpretable time series forecasting', in *ICLR*, (2020).

[28] Jacopo Pavan, Francesco Prendin, Lorenzo Meneghetti, Giacomo Cappon, Giovanni Sparacino, Andrea Facchinetti, and Simone Del Favero, 'Personalized machine learning algorithm based on shallow network and error imputation module for an improved blood glucose prediction', in *this volume*, (August 2020).

[29] Harry Rubin-Falcone, Ian Fox, and Jenna Wiens, 'Deep residual time-series forecasting: Application to blood glucose prediction', in *this volume*, (August 2020).

[30] Xiaoyu Sun, Mudassir Rashid, Mert Sevil, Nicole Hobbs, Rachel Brandt, Mohammad Reza Askari, Andrew Shahidehpour, and Ali Cinar, 'Prediction of blood glucose levels for people with type 1 diabetes using latent-variable-based model', in *this volume*, (August 2020).

[31] Tao Yang, Ruikun Wu, Rui Tao, Shuang Wen, Ning Ma, Yuhang Zhao, Xia Yu, and Hongru Li, 'Multi-scale long short-term memory network with multi-lag structure for blood glucose prediction', in *this volume*, (August 2020).

[32] Taiyu Zhu, Xi Yao, Kezhi Li, Pau Herrero, and Pantelis Georgiou, 'Blood glucose prediction for type 1 diabetes using generative adversarial networks', in *this volume*, (August 2020).

KDH Workshop Co-Chairs

- Kerstin Bach, Norwegian University of Science and Technology
- Cindy Marling, Ohio University
- Nirmalie Wiratunga, The Robert Gordon University

BGLP Challenge Co-Chairs

- Razvan Bunescu, Ohio University
- Cindy Marling, Ohio University

Steering Committee

- Kerstin Bach, Norwegian University of Science and Technology
- Sadid Hasan, Philips Research North America
- Zina Ibrahim, King's College London
- Cindy Marling, Ohio University
- Jonathan Rubin, Philips Research North America
- Nirmalie Wiratunga, The Robert Gordon University
- Honghan Wu, University of Edinburgh

Program Committee

- Isabelle Bichindaritz, State University of New York at Oswego
- Ali Cinar, Illinois Institute of Technology
- José Manuel Colmenar, Universidad Rey Juan Carlos
- Alexandra Constantin, Bigfoot Biomedical
- Iván Contreras, Universidad Complutense de Madrid
- Vivek V Datla, Philips Research North America
- Antonio Della Cioppa, University of Salerno
- Spiros Denaxas, University College London
- Franck Dernoncourt, Massachusetts Institute of Technology
- Andrea Facchinetti, University of Padova
- Ian Fox, University of Michigan
- Pau Herrero, Imperial College London
- J. Ignacio Hidalgo, Universidad Complutense de Madrid
- Fernando Koch, IBM Global Services
- Kezhi Li, University College London
- Stewart Massie, Robert Gordon University
- Stefania Montani, University of Piemonte Oriental
- Stavroula Mougiakakou, University of Bern
- Tristan Naumann, Microsoft
- Alexander Schliep, Gothenburg University
- Thomas Searle, Kings College London
- Giovanni Sparacino, University of Padova
- Shawn Stapleton, Philips Research North America
- Lukas Stappen, University of Augsburg
- Josep Vehi, University of Girona
- Anjana Wijekoon, The Robert Gordon University