Integrating Explainable Machine Learning and Predictive Process Monitoring

Williams Rizzi[1,2]

[1]Fondazione Bruno Kessler, Trento, Italy
[2]Free University of Bozen-Bolzano, Bolzano, Italy

1. Introduction

My PhD is focused on improving Predictive Process Monitoring (PPM) in several directions. The direction I am currently investigating is the improvement of PPM by exploiting explainable machine learning techniques.

Predictive Process Monitoring [1] is a research topic aiming at developing techniques that event logs extracted from information systems to predict how ongoing process executions will unfold up to their completion.

Explainable Machine Learning [2] is a research topic within the field of Explainable Artificial Intelligence (XAI), aiming at explaining why a given predictive model behaves in a certain way or in general how it will behave. Explainability techniques have been proven to be mature to be used, yet they have not been evaluated in-depth in the PPM scenario.

In this document, I focus on incorporating explainability techniques in PPM. In the PPM scenario we decided to identify three different actors to benefit from our work: (i) the business analyst; (ii) the machine; and (iii) the research scientist.

First, the document introduces the challenges of the work (Section 2), the approaches that have been/are going to be used for facing these challenges (Section 3), the results obtained so far (Section 4), and finally the position w.r.t. the State-of-the-art (Section 5).

2. Challenges

The three actors we identified have different understandings of the PPM scenario and different needs, this does not allow us to devise a solution that fits them all. To give the benefit of explainability techniques to the three aforementioned actors we focus on three challenges: (i) allowing business analysts to understand predictions and predictive models; (ii) allowing machines to enhance predictive models using the information gathered by the explainer as feedback; and (iii) allowing research scientists to leverage, adapt and build on top of explainability techniques for PPM.

Proceedings of the Demonstration & Resources Track, Best BPM Dissertation Award, and Doctoral Consortium at BPM 2021 co-located with the 19th International Conference on Business Process Management, BPM 2021, Rome, Italy, September 6-10, 2021
wrizzi@fbk.eu (W. Rizzi)
0000-0002-7318-6833 (W. Rizzi)

The first challenge is using explainability techniques to allow business analysts to understand predictions and predictive models and to support them in other tasks, like decision making. The classical output of the explainability techniques have been tested in the general use-case of predictive model usage but not in the PPM scenario. What is missing is an investigation of whether the business analysts can be empowered with the same techniques. Empowering business analysts with the explanations for the predictive model could result in getting the trust of such users as well as in supporting them in making more informed decisions.

The second challenge is to allow machines to enhance the predictive model, by making use of the explanation as feedback. Since the explanation allows us to understand why the predictive model is behaving in a given way, we can use this explanation to characterise the wrong predictions. Understanding how the predictive model produces a wrong prediction would allow the machine to perform an automatic prevention, or at least correction, of the wrong predictions.

The third challenge is supporting research scientists with a representative set of explainability techniques for PPM that they can leverage and adapt to new problems. Up to now indeed, there is no code or API support for predictive model or prediction explanation out-of-the-box for PPM. Supporting the explanation of PPM models and data could be useful as a starting point for the research scientists who can build new solutions on top of them.

3. Addressing the Challenges

To address the aforementioned challenges we followed the methodology described in [3]. Each of the aforementioned challenges is addressed with the delivery of different artifacts and their respective validation. In particular, the used/planned approaches are: (i) the development/adoption of visual representation of the explainability techniques' output and a user evaluation to understand how to empower the business analysts with the explainability techniques' output, (ii) a technique to enhance predictive models through explanations, and (iii) the integration of a representative set of explainability techniques for PPM into a framework as easy-to-use APIs.

The first challenge is addressed with the development/adoption of several different visualisations of explanations and a user evaluation. In the user evaluation, we investigate whether explanations for PPM are understandable by business analysts. We then investigate whether business analysts are capable to exploit the information contained in the explanation to perform other tasks; e.g., decision making.

The second challenge is addressed developing and evaluating an algorithm that aims at understanding the reasons why a prediction is wrong and leveraging this information for improving the accuracy of the model. The algorithm mines explanation patterns from sets of traces, that aim at characterising the wrong predictions and use these patterns to lower down the relevance of the same patterns in the training set. To this aim, the returned predictions, each referring to an incomplete trace, are classified based on the value of the prediction and on the value for the ground truth. For instance, in the case of binary predictions, the prediction related to each trace is classified in terms of one of the confusion matrix quadrants, i.e., true positives, true negatives, false positives, and false negatives. Patterns of explanations can then be identified over each class of traces. The machine can then leverage patterns characterising the wrong predictions by acting upon the training set reducing their relevance. Finally, the predictive model can be re-trained with the new training set possibly presenting less misleading examples than those causing the initial predictive model to learn patterns determining wrong predictions.

The third challenge is addressed by providing a representative set of explainability techniques although most of the PPM solutions exploit common machine learning models so that explainability techniques can be easily applied to the PPM scenarios, the main point of attention required for adapting explainability techniques to the PPM scenario is related to how the PPM data are encoded/decoded, due to the particular type of data used in this domain.

4. Results

Tackling the three aforementioned challenges produced the following results.

Results for the first challenge. We do not have conclusive results for this challenge, we are carrying out the user-evaluation. To evaluate the quality of the explainability solutions for PPM we aim at evaluating: (i) if the user understands the explanations; and (ii) if the explanation can influence positively the work of the business analyst. To evaluate *(i)* we give the user several different explanations, and we gather its insights on what is understandable. To evaluate *(ii)* we ask the user to carry out several decision-making tasks with the explainability information we provide.

Results for the second challenge. This is the challenge for which we have the most tangible results. In the paper [4] we built a working solution that exploits feature-importance explanations to explain why a predictive model is wrong and eventually uses these feature-importance explanations to improve its accuracy. Given a trained predictive model, our solution computes the explanations for each of the traces of the validation set. Depending on the received predictions and on the actual labels, the traces are assigned to a quadrant of the confusion matrix. The traces contained in the quadrants of the correct and wrong predictions are then filtered keeping only the important features, using the feature-importance values received by the explainability technique. The filtered traces are then fed to a pattern miner, which retrieves the patterns characterising the correct and wrong predictions. We identify the occurrence of the patterns characterising the wrong predictions in the initial training set. The features identified by the patterns characterising the wrong predictions inside the identified set of traces are then shuffled to destroy the occurrence of these patterns, we then re-train the model and test it. This results in a significant improvement of the accuracy both on synthetic and real datasets. This solution currently works for the binary prediction problem and we are extending it to support the multiclass problem. The designed algorithm relies on Random Forest [5] and LIME [6], we identified the work in [7] as compatible with our approach and we are currently working with the authors on integrating our approaches to extend the significance of our respective contributions.

Results for the third challenge. To support the research scientists, we are continuing the development of Nirdizati [8]. We identified the works in (i) Partial Dependence Plot (PDP) [9], (ii) Individual Conditional Expectations (ICE) [10], (iii) Local Interpretable Model Explanations (LIME) [6], (iv) Shapley values for Explanation (SHAP) [11], (v) Skater [12], and (vi) Anchors [13] as suitable and representative of the state-of-the-art for explainability. We integrated them in Nirdizati, built the explanation module, and provided them as APIs.

5. Related Works

The body of previous work related to these challenges is the one that concerns: (i) PPM in general, (ii) the application of explainability techniques in PPM, and (iii) the development of PPM frameworks.

For what concerns the body of work related to PPM in general we can identify three streams of work: (i) outcome-oriented prediction [14, 1, 15, 16], (ii) numeric/remaining time prediction [17, 18, 19], and (iii) next activity prediction [20, 21, 22, 23, 24]. These works differ from the work in [4] since they build the model once and for all and do not try to actively enhance the performance of the predictive model, or give any hint to the user on what moves the predictive model.

For what concerns the body of work related to the application of explainability techniques in PPM we identified [25, 7, 26, 27, 28]. We can divide them in those using a post-hoc explainability technique [25, 7, 26, 27], and the one using an ante-hoc explainability technique [25]. All these works make next activity prediction, [25, 7, 26, 27] with deep neural networks, whereas [28] with Bayesian Networks, and they all aim at providing the user with an understanding of what moved the predictive model towards a prediction. None of these works: (i) empirically evaluates the explanation delivered to the user, and (ii) actively enhance the accuracy of the predictive model.

For what concerns the body of work related to the development of PPM frameworks we identified, some plug-ins in ProM [29], for instance [30, 18, 1, 31] and one plug-in in Apromore [32], which is [33]. They differ from our work on integrating explainability techniques in PPM frameworks because they do not support explainability techniques by design.

Acknowledgements

I want to thank my supervisors, Chiara Ghidini (FBK), Chiara Di Francescomarino (FBK), and Fabrizio Maria Maggi (UNIBZ), for many hours of discussion and feedback regarding the research topic and for their help to formulate the thesis subject.

References

[1] F. M. Maggi, C. Di Francescomarino, M. Dumas, C. Ghidini, Predictive monitoring of business processes, in: M. Jarke, J. Mylopoulos, C. Quix, C. Rolland, Y. Manolopoulos, H. Mouratidis, J. Horkoff (Eds.), Advanced Information Systems Engineering - 26th International Conference, CAiSE 2014, Thessaloniki, Greece, June 16-20, 2014. Proceedings, volume 8484 of *Lecture Notes in Computer Science*, Springer, 2014, pp. 457–472. doi:10.1007/978-3-319-07881-6_31.

[2] S. T. Mueller, R. R. Hoffman, W. J. Clancey, A. Emrey, G. Klein, Explanation in human-ai systems: A literature meta-review, synopsis of key ideas and publications, and bibliography for explainable AI, CoRR abs/1902.01876 (2019). URL: http://arxiv.org/abs/1902.01876. arXiv:1902.01876.

[3] J. vom Brocke, A. Hevner, A. Maedche, Introduction to design science research, in: Design Science Research. Cases, Springer, 2020.

[4] W. Rizzi, C. Di Francescomarino, F. M. Maggi, Explainability in predictive process monitoring: When understanding helps improving, in: D. Fahland, C. Ghidini, J. Becker, M. Dumas (Eds.), Business Process Management Forum - BPM Forum 2020, Seville, Spain, September 13-18, 2020, Proceedings, volume 392 of *Lecture Notes in Business Information Processing*, Springer, 2020, pp. 141–158. doi:10.1007/978-3-030-58638-6_9.
[5] T. K. Ho, Random decision forests, in: Third International Conference on Document Analysis and Recognition, ICDAR 1995, August 14 - 15, 1995, Montreal, Canada. Volume I, IEEE Computer Society, 1995, pp. 278–282. doi:10.1109/ICDAR.1995.598994.
[6] M. T. Ribeiro, S. Singh, C. Guestrin, "why should I trust you?": Explaining the predictions of any classifier, CoRR abs/1602.04938 (2016). arXiv:1602.04938.
[7] S. Weinzierl, S. Zilker, J. Brunk, K. Revoredo, M. Matzner, J. Becker, XNAP: making lstm-based next activity predictions explainable by using LRP, in: A. del-Río-Ortega, H. Leopold, F. M. Santoro (Eds.), Business Process Management Workshops - BPM 2020 International Workshops, Seville, Spain, September 13-18, 2020, Revised Selected Papers, volume 397 of *Lecture Notes in Business Information Processing*, Springer, 2020, pp. 129–141. doi:10.1007/978-3-030-66498-5_10.
[8] W. Rizzi, L. Simonetto, C. Di Francescomarino, C. Ghidini, T. Kasekamp, F. M. Maggi, Nirdizati 2.0: New features and redesigned backend, in: B. Depaire, J. D. Smedt, M. Dumas, D. Fahland, A. Kumar, H. Leopold, M. Reichert, S. Rinderle-Ma, S. Schulte, S. Seidel, W. M. P. van der Aalst (Eds.), Proceedings of the Dissertation Award, Doctoral Consortium, and Demonstration Track at BPM 2019 co-located with 17th International Conference on Business Process Management, BPM 2019, Vienna, Austria, September 1-6, 2019, volume 2420 of *CEUR Workshop Proceedings*, CEUR-WS.org, 2019, pp. 154–158. URL: http://ceur-ws.org/Vol-2420/paperDT8.pdf.
[9] J. H. Friedman, Greedy function approximation: a gradient boosting machine, Annals of statistics (2001) 1189–1232.
[10] A. Goldstein, A. Kapelner, J. Bleich, E. Pitkin, Peeking inside the black box: Visualizing statistical learning with plots of individual conditional expectation, Journal of Computational and Graphical Statistics 24 (2015) 44–65.
[11] S. M. Lundberg, S. Lee, A unified approach to interpreting model predictions, in: I. Guyon, U. von Luxburg, S. Bengio, H. M. Wallach, R. Fergus, S. V. N. Vishwanathan, R. Garnett (Eds.), Advances in Neural Information Processing Systems 30: Annual Conference on Neural Information Processing Systems 2017, December 4-9, 2017, Long Beach, CA, USA, 2017, pp. 4765–4774.
[12] A. Kramer, et al., Skater, https://github.com/oracle/Skater/, 2018.
[13] M. T. Ribeiro, S. Singh, C. Guestrin, Anchors: High-precision model-agnostic explanations, in: S. A. McIlraith, K. Q. Weinberger (Eds.), Proceedings of the Thirty-Second AAAI Conference on Artificial Intelligence, (AAAI-18), the 30th innovative Applications of Artificial Intelligence (IAAI-18), and the 8th AAAI Symposium on Educational Advances in Artificial Intelligence (EAAI-18), New Orleans, Louisiana, USA, February 2-7, 2018, AAAI Press, 2018, pp. 1527–1535. URL: https://www.aaai.org/ocs/index.php/AAAI/AAAI18/paper/view/16982.
[14] I. Teinemaa, M. Dumas, M. La Rosa, F. M. Maggi, Outcome-oriented predictive process monitoring: Review and benchmark, TKDD 13 (2019) 17:1–17:57. doi:10.1145/3301300.

[15] C. Di Francescomarino, M. Dumas, F. M. Maggi, I. Teinemaa, Clustering-based predictive process monitoring, IEEE Trans. Services Computing 12 (2019) 896–909. doi:10.1109/TSC.2016.2645153.
[16] A. Leontjeva, R. Conforti, C. Di Francescomarino, M. Dumas, F. M. Maggi, Complex symbolic sequence encodings for predictive monitoring of business processes, in: H. R. Motahari-Nezhad, J. Recker, M. Weidlich (Eds.), Business Process Management - 13th International Conference, BPM 2015, Innsbruck, Austria, August 31 - September 3, 2015, Proceedings, volume 9253 of *Lecture Notes in Computer Science*, Springer, 2015, pp. 297–313. doi:10.1007/978-3-319-23063-4_21.
[17] W. M. P. van der Aalst, M. H. Schonenberg, M. Song, Time prediction based on process mining, Inf. Syst. 36 (2011) 450–475. doi:10.1016/j.is.2010.09.001.
[18] F. Folino, M. Guarascio, L. Pontieri, Discovering context-aware models for predicting business process performances, in: R. Meersman, H. Panetto, T. S. Dillon, S. Rinderle-Ma, P. Dadam, X. Zhou, S. Pearson, A. Ferscha, S. Bergamaschi, I. F. Cruz (Eds.), On the Move to Meaningful Internet Systems: OTM 2012, Confederated International Conferences: CoopIS, DOA-SVI, and ODBASE 2012, Rome, Italy, September 10-14, 2012. Proceedings, Part I, volume 7565 of *Lecture Notes in Computer Science*, Springer, 2012, pp. 287–304. doi:10.1007/978-3-642-33606-5_18.
[19] A. Rogge-Solti, M. Weske, Prediction of remaining service execution time using stochastic petri nets with arbitrary firing delays, in: S. Basu, C. Pautasso, L. Zhang, X. Fu (Eds.), Service-Oriented Computing - 11th International Conference, ICSOC 2013, Berlin, Germany, December 2-5, 2013, Proceedings, volume 8274 of *Lecture Notes in Computer Science*, Springer, 2013, pp. 389–403.
[20] N. Tax, I. Verenich, M. La Rosa, M. Dumas, Predictive business process monitoring with LSTM neural networks, in: Proceedings of the International Conference on Advanced Information Systems Engineering (CAiSE), 2017, pp. 477–492.
[21] C. Di Francescomarino, C. Ghidini, F. M. Maggi, G. Petrucci, A. Yeshchenko, An eye into the future: Leveraging a-priori knowledge in predictive business process monitoring, in: BPM, 2017, pp. 252–268.
[22] M. Camargo, M. Dumas, O. G. Rojas, Learning accurate LSTM models of business processes, in: T. T. Hildebrandt, B. F. van Dongen, M. Röglinger, J. Mendling (Eds.), Business Process Management - 17th International Conference, BPM 2019, Vienna, Austria, September 1-6, 2019, Proceedings, volume 11675 of *Lecture Notes in Computer Science*, Springer, 2019, pp. 286–302. doi:10.1007/978-3-030-26619-6_19.
[23] J. Brunk, J. Stottmeister, S. Weinzierl, M. Matzner, J. Becker, Exploring the effect of context information on deep learning business process predictions, Journal of Decision Systems (2020) 1–16.
[24] F. Taymouri, M. La Rosa, S. M. Erfani, Z. D. Bozorgi, I. Verenich, Predictive business process monitoring via generative adversarial nets: The case of next event prediction, in: D. Fahland, C. Ghidini, J. Becker, M. Dumas (Eds.), Business Process Management - 18th International Conference, BPM 2020, Seville, Spain, September 13-18, 2020, Proceedings, volume 12168 of *Lecture Notes in Computer Science*, Springer, 2020, pp. 237–256. doi:10.1007/978-3-030-58666-9_14.
[25] R. Galanti, B. Coma-Puig, M. de Leoni, J. Carmona, N. Navarin, Explainable predictive

process monitoring, in: B. F. van Dongen, M. Montali, M. T. Wynn (Eds.), 2nd International Conference on Process Mining, ICPM 2020, Padua, Italy, October 4-9, 2020, IEEE, 2020, pp. 1–8. doi:10.1109/ICPM49681.2020.00012.

[26] R. Sindhgatta, C. Moreira, C. Ouyang, A. Barros, Exploring interpretable predictive models for business processes, in: D. Fahland, C. Ghidini, J. Becker, M. Dumas (Eds.), Business Process Management - 18th International Conference, BPM 2020, Seville, Spain, September 13-18, 2020, Proceedings, volume 12168 of *Lecture Notes in Computer Science*, Springer, 2020, pp. 257–272. doi:10.1007/978-3-030-58666-9_15.

[27] J. Rehse, N. Mehdiyev, P. Fettke, Towards explainable process predictions for industry 4.0 in the dfki-smart-lego-factory, Künstliche Intell. 33 (2019) 181–187. doi:10.1007/s13218-019-00586-1.

[28] S. Pauwels, T. Calders, Bayesian network based predictions of business processes, in: D. Fahland, C. Ghidini, J. Becker, M. Dumas (Eds.), Business Process Management Forum - BPM Forum 2020, Seville, Spain, September 13-18, 2020, Proceedings, volume 392 of *Lecture Notes in Business Information Processing*, Springer, 2020, pp. 159–175. doi:10.1007/978-3-030-58638-6_10.

[29] B. F. van Dongen, A. K. A. de Medeiros, H. M. W. Verbeek, A. J. M. M. Weijters, W. M. P. van der Aalst, The prom framework: A new era in process mining tool support, in: G. Ciardo, P. Darondeau (Eds.), Applications and Theory of Petri Nets 2005, 26th International Conference, ICATPN 2005, Miami, USA, June 20-25, 2005, Proceedings, volume 3536 of *Lecture Notes in Computer Science*, Springer, 2005, pp. 444–454. doi:10.1007/11494744_25.

[30] M. de Leoni, W. M. P. van der Aalst, M. Dees, A general framework for correlating business process characteristics, in: S. W. Sadiq, P. Soffer, H. Völzer (Eds.), Business Process Management - 12th International Conference, BPM 2014, Haifa, Israel, September 7-11, 2014. Proceedings, volume 8659 of *Lecture Notes in Computer Science*, Springer, 2014, pp. 250–266. doi:10.1007/978-3-319-10172-9_16.

[31] M. Federici, W. Rizzi, C. Di Francescomarino, M. Dumas, C. Ghidini, F. M. Maggi, I. Teinemaa, A prom operational support provider for predictive monitoring of business processes, in: F. Daniel, S. Zugal (Eds.), Proceedings of the BPM Demo Session 2015 Co-located with the 13th International Conference on Business Process Management (BPM 2015), Innsbruck, Austria, September 2, 2015, volume 1418 of *CEUR Workshop Proceedings*, CEUR-WS.org, 2015, pp. 1–5. URL: http://ceur-ws.org/Vol-1418/paper1.pdf.

[32] M. La Rosa, H. A. Reijers, W. M. P. van der Aalst, R. M. Dijkman, J. Mendling, M. Dumas, L. García-Bañuelos, APROMORE: an advanced process model repository, Expert Syst. Appl. 38 (2011) 7029–7040. doi:10.1016/j.eswa.2010.12.012.

[33] I. Verenich, S. Mõskovski, S. Raboczi, M. Dumas, M. La Rosa, F. M. Maggi, Predictive process monitoring in apromore, in: J. Mendling, H. Mouratidis (Eds.), Information Systems in the Big Data Era - CAiSE Forum 2018, Tallinn, Estonia, June 11-15, 2018, Proceedings, volume 317 of *Lecture Notes in Business Information Processing*, Springer, 2018, pp. 244–253. doi:10.1007/978-3-319-92901-9_21.