

KnowledgePit Meets BrightBox: A Step Toward Insightful Investigation of the Results of Data Science Competitions

Andrzej Janusz*[†], Dominik Ślęzak*[†]

*Institute of Informatics, University of Warsaw, Warsaw, Poland

[†]QED Software, Warsaw, Poland

Email: {firstname.lastname}@qed.pl

Abstract—We discuss the benefits of integrating the KnowledgePit data science competition platform with the BrightBox technology aimed at diagnostics of machine learning models embedded within complex software systems. We briefly recall the history of international challenges held at KnowledgePit and we also discuss in what sense such technologies as BrightBox can be helpful during the post-challenge analysis. In particular, we show how to combine solutions submitted by the competition participants in order to obtain even more accurate predictions. The discussed functionalities are of significant importance for the sponsors and organizers of data science / machine learning online contests because they support adoption of submissions while designing ultimate solutions of real-world problems.

Index Terms—Data science competitions; machine learning; model stacking; KnowledgePit platform; BrightBox technology

I. INTRODUCTION

KNOWLEDGE Pit¹ is an online platform for organizing data science / machine learning (ML) challenges. Its architecture was first presented in [1] and since then, it has been improving continually. Currently, KnowledgePit puts together the functionalities of a typical competition platform – such as Kaggle² – with additional tools that make it possible for the competition sponsors and organizers to investigate the submitted solutions with respect to their *true* usefulness in the corresponding real-life decision problems. These tools are available thanks to integrating KnowledgePit with BrightBox – the technology developed by QED Software³ for the purpose of assessing the decision models basing on the analysis of mistakes that they are making [2]. In this paper, we discuss one of such functionalities – designed at the border of KnowledgePit and BrightBox – which lets us create better models by mixing solutions acquired from the competition participants.

The paper is organized as follows: In Section II, we recall the main ideas behind KnowledgePit and, as an illustration, we report the history of the KnowledgePit contests held in cooperation with the FedCSIS conference series. Analogously, Section III introduces the main ideas behind BrightBox, with a special emphasis on its contributions into the KnowledgePit’s

functionality. Section IV refers to the data science challenge which was associated with this year’s FedCSIS conference⁴. Besides describing the competition itself, we include here some KnowledgePit-supported visualizations that can be helpful for sponsors and organizers. In Section V, we explain our aforementioned idea of mixing the competition solutions and report the experimental results obtained for the challenge outlined in Section IV. Section VI concludes the paper.

II. THE HISTORY OF KNOWLEDGE PIT

The platforms such as Kaggle attract thousands of data scientists to participate in challenges aiming at solving real-life problems. Such challenges not only address specific problems but often facilitate innovative applications of ML algorithms. On the one hand, they are appealing to those for whom competitive challenges can be a source of new interesting research topics. They can also be an attractive addition to academic courses for students who are interested in practical applications. On the other hand, setting up a public data science competition is a form of outsourcing a given task to the community [3]. It can be beneficial to the sponsors and organizers who set up the contest, as it is an inexpensive approach to solve the problem that they are after [4].

Accordingly, it should not be surprising that the scope of our own platform – KnowledgePit – shifted during the years from organizing smaller, mostly student-focused challenges and projects to international data science competitions. Although KnowledgePit still hosts several student competitions for ML-related university courses every year, the most prestigious events are those prepared for big industry clients and partners, in association with international conferences [5], [6].

One may say that our competitions grew together with recognition of the FedCSIS conferences. Together with the one reported in Section IV [7], there have been already nine challenges held at KnowledgePit in cooperation with FedCSIS. The series started in 2014 with the *AAIA’14 Data Mining Competition: Key Risk Factors for Polish State Fire Service* [8]. Other competition topics included the recognition of firefighters’ activities based on inertial sensor readings [9], predicting seismic activity in coal mines [10], predicting video

This research was co-funded by the Polish National Centre for Research and Development in frame of project MAZOWSZE/0198/19.

¹<https://knowledgepit.ai/>

²<https://www.kaggle.com>

³<https://qed.pl/project/brightbox>

⁴<https://knowledgepit.ai/fedcsis-2022-challenge/>

game winners based on game logs [11], marking hair follicles on microscopic images⁵, predicting win-rates of custom card decks in collectible card games [12], [13], and predicting typical patterns in network device workloads [14]. All of these competitions were highly successful. With more than 1,300 participating teams and several thousands of submitted solutions, they significantly contributed to solving important real-life challenges. They also provided us with a comprehensive survey on the state-of-the-art ML approaches in the related fields, such as time series forecasting [15], feature extraction [16], as well as prediction model ensembling [17].

III. KNOWLEDGEPIT MEETS BRIGHTBOX

During its journey, KnowledgePit had to evolve to fit the needs of our industrial partners. One of the most significant needs has been related to the post-competition analysis of the submitted solutions. With this regard, it was possible to meet the industry expectations thanks to integrating KnowledgePit with the aforementioned BrightBox technology.

As highlighted in Section I, BrightBox is a software technology which assesses decision models (aimed at classification, regression, prediction, etc.) basing on their mistakes (i.e. differences between their outputs and the observed ground truth). Its main application field is in diagnostics of ML models deployed within complex systems [18]. Its methods have deep roots in the theory of rough sets [19]. A diagnosed model is approximated by the rough-set-based surrogate models – the ensembles of so-called approximate decision reducts [20]. Then, particular cases are investigated by looking at their neighborhoods – the groups of other cases that are classified similarly by the surrogate reducts (i.e. objects that fit the same rules induced by reducts [21]). If a mistake that happened for a given case often repeats in its neighborhood, then it is likely that the diagnosed model was not trained correctly for such objects from the beginning. As another example, if the neighborhood is almost empty (i.e.: the diagnosed object seems to be classified by different rules than objects observed before), then BrightBox can conclude that the model seems to find itself in a new situation. Such hints are useful from an operational viewpoint when it comes to rebuilding ML models as parts of the aforementioned complex systems.

Although it was a non-trivial effort to adapt the default settings of BrightBox to the specifics of a data science competition platform (for instance, we had to address different scenarios of the access to the diagnosed models' behavior on the training data), it enabled us to extend basic KnowledgePit's functionalities (such as the analysis of trends in the quality scores or survey-based summaries of commonly applied ML techniques) with in-depth diagnostics of individual submissions. Since BrightBox does not require a direct access to the diagnosed models, it can be applied to construct the above-discussed surrogate models that approximate submitted solutions and allow reasoning about their properties. For example, it allows to approximate feature importance coefficients of

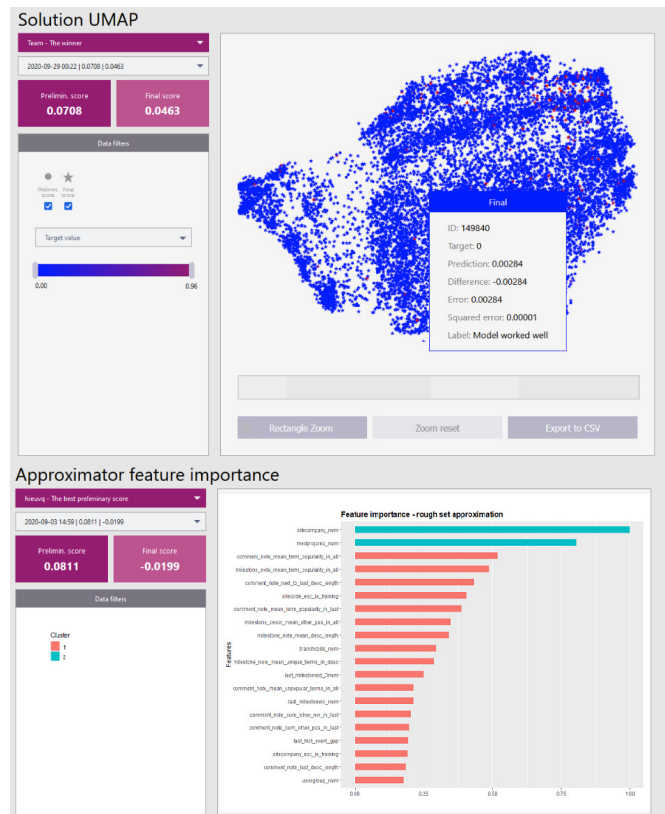


Fig. 1: Exemplary visualizations of one of solutions from challenge [5]. Points on the upper plot correspond to cases from the test data and their color reflects prediction errors. The lower plot shows approximated (using BrightBox) feature importance for the model used to create the analyzed submission. An experimental evaluation showed that the Spearman correlation between the approximated feature importance values and the actual values estimated for the model was ≈ 0.7 .

models used to create the submissions (see Figure 1). It can also provide insightful information on types of errors committed by models, and similarities between solutions submitted by competition participants. Prototype implementations of some selected functionalities provided by the BrightBox technology have been already tested in our previous competitions [5], [22].

IV. FEDCSIS 2022 CHALLENGE

As mentioned in Section II, FedCSIS 2022 hosted the ninth KnowledgePit competition associated with the FedCSIS series. We use this competition as an illustration how KnowledgePit works, as well as a prerequisite for the discussion in Section V. We refer to [7] for more details about this particular contest. The best solutions submitted by its participants are described in [23] (1st place), [24] (2nd), [25] (3rd), and [26] (4th).

The task was to predict the execution costs of so-called forwarding contracts. The data about contracts was provided by the competition sponsor – a company that develops decision

⁵<https://knowledgepit.ai/esensei-challenge/>

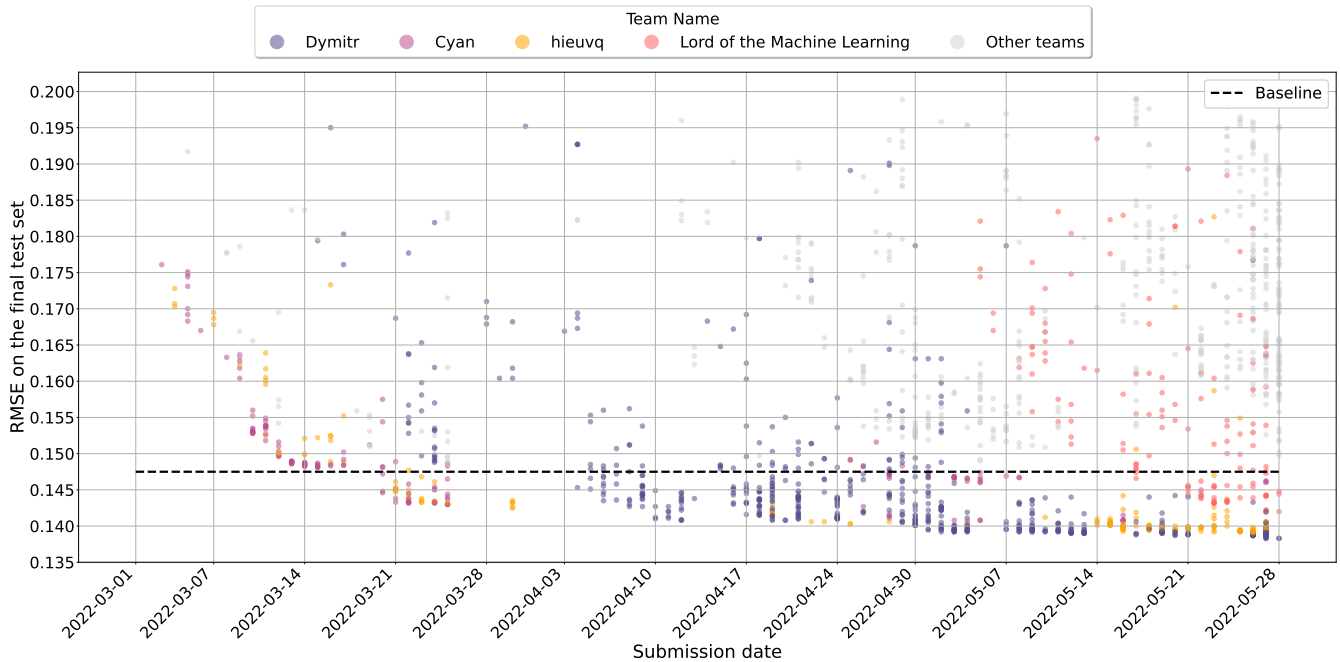


Fig. 2: The final scores of all solutions that reached less than 0.2 RMSE. The most successful teams are specially highlighted.

TABLE I: Top results of the FedCSIS 2022 Challenge [7]. Both preliminary and final scores reflect the RMSE measure.

rank	team name	preliminary	final	#submissions
1	Dymitr [23]	0.1398	0.1383	619
2	Cyan [24]	0.1402	0.1391	181
3	hieuvq [25]	0.1396	0.1407	159
4	Lord of the ML [26]	0.1434	0.1420	147
5	baseline [7]	0.1491	0.1475	–

TABLE II: The RMSE results of the solution stacking methods on the final test set. The simple solution averaging and the weighted solution averaging are included as a reference.

ensemble	final	#solutions
<i>average - best solution of top 5 teams</i>	0.1361	5
<i>weighted average - best solution of top 5 teams</i>	0.1360	5
<i>average - best solution of top 8 teams</i>	0.1366	8
<i>weighted average - best solution of top 8 teams</i>	0.1367	8
<i>LASSO reg. (opt. λ) best solution of each team</i>	0.1344	8
<i>average - top 28 solutions</i>	0.1358	28
<i>weighted average - top 28 solutions</i>	0.1358	28
<i>LASSO reg. (opt. λ) all solutions</i>	0.1339	28

support and optimization systems in the transportation, spedition, and logistics areas. That company was highly interested in a deeper analysis of the submitted solutions with respect to their potential deployment within the designed systems.

For the evaluation purposes, the data was divided into the training set (330,055 historical contracts) and test set (72,452 newer contracts). The data was carefully anonymized, but it

was done in such a way that its analytical value is not lost (see also our other competitions [5]). Given a regression nature of the considered prediction task, we selected one of typical evaluation measures – the root mean square error (RMSE). However, let us note that sometimes specifying a measure that truly reflects a real-life decision problem is not easy [27].

The competition was conducted in a standard way including: (i) the online preliminary evaluation on an unknown subset of the test data (the participants submit solutions for the full test set but preliminary evaluation is done on a subset which is unknown to them), (ii) the associated public leaderboard (only preliminary results are shown during the competition), and (iii) the final evaluation on the full test set (which is calculated after the competition is closed, after the participants select their final solutions to be evaluated, and only if they submit to KnowledgePit the reports that describe their solutions).

Table I displays the top results. Let us note that the preliminary and final rankings are different. Actually, the solution at the 3rd final place would be the best one if only preliminary scores were taken into account. Such cases are specially worth investigating with the usage of the BrightBox technology.

Table I shows also the baseline solution – the model that we prepared by ourselves prior to the contest’s beginning [7]. In this challenge, only four teams exceeded the baseline score (though sometimes it may be even less, see e.g. [14]). Figure 2 displays the history of all submissions of those teams.

V. MIXING THE SOLUTIONS

Due to a large number of submitted solutions and the fact that most of the best-performing teams used some advanced ensembling methods to compute their final predictions [23],

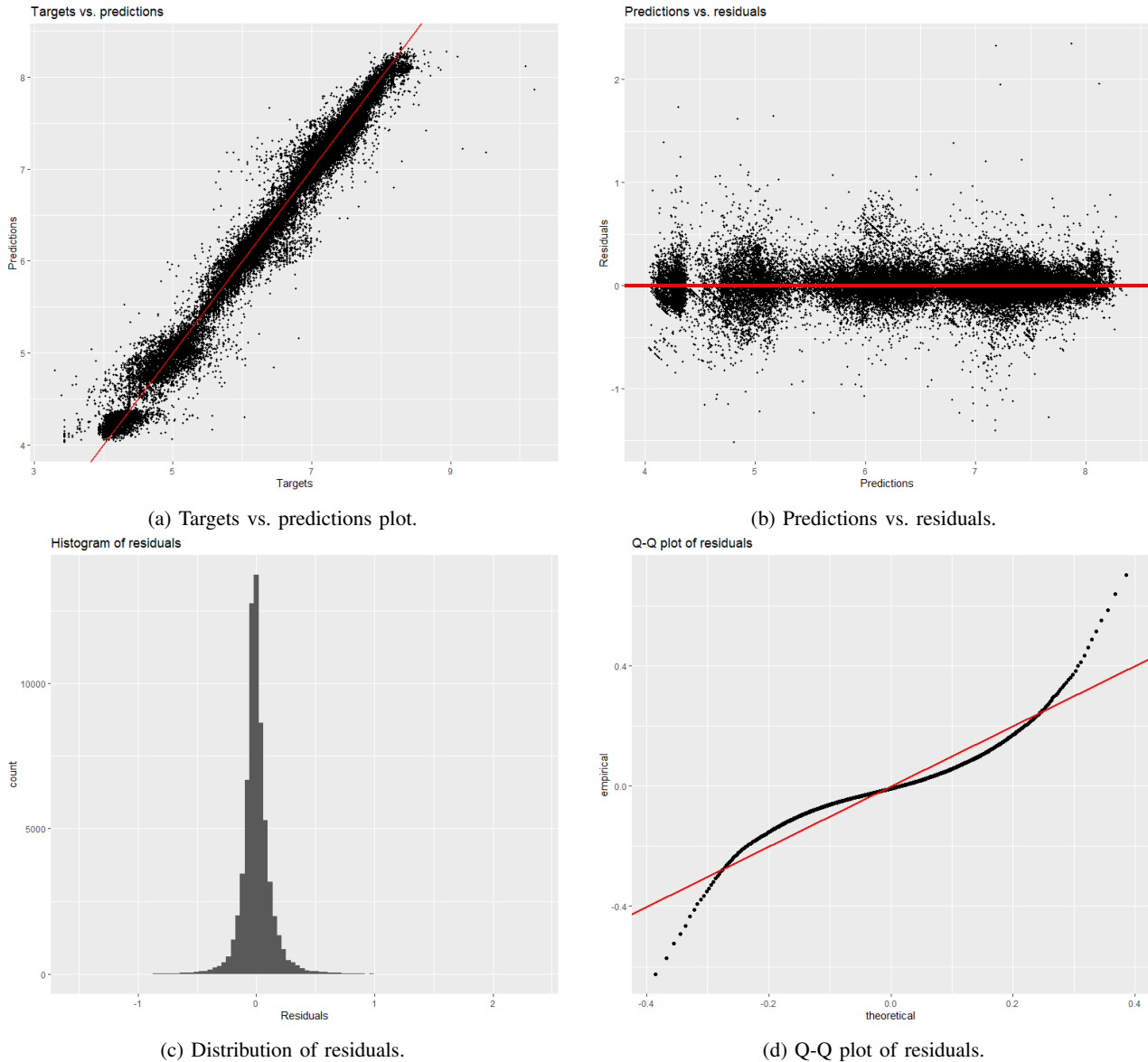


Fig. 3: Visual quality analysis of the ensemble model obtained using the considered solution mixing methodology. The top-left plot shows the relation between ground truths and predictions for test cases. The top-right scatter plot visualizes residuals with regard to predictions. The bottom-left plot shows the distribution of residuals on test cases, and the bottom-right plot shows the relation between theoretical quantiles of the Gaussian distribution and the empirical quantiles of the residuals.

[24], [25], [26], we decided to check how accurate the predictions could be if we mix submissions from different teams. In this way, not only could we verify which solutions are the most influential and potentially worth deploying, but also we could get a better insight into where the limit to prediction quality is when working with this type of the data.

To create the ensemble of submissions, we used the post-competition data analytics tools provided by KnowledgePit thanks to its integration with BrightBox (see also [22]). We compared the results of the LASSO regression-based model stacking [28] to simple model averaging methods. We also

investigated two ensembling scenarios. In the first one, we used all correctly formatted submissions. In the second one, we restrained the submissions to those with the best score in the preliminary evaluation from each team. In both cases, we trained the logistic regression model with LASSO regularization on predictions for the competition's preliminary evaluation set. Then, we verified the performance of each ensemble on the same final test set as all individual submissions.

We tuned the regularization parameter λ of the LASSO regression using the cross-validation technique, following the

guidelines taken from the `glmnet` package⁶. We chose λ for which the validation loss was lowest. Table II shows the results for the obtained ensembles. As a reference, we include the results obtained using simple averaging of the best solutions from top 5 teams (solutions [23], [24], [25], [26] and the baseline model), and the results of weighted averaging in which the weights correspond to the preliminary scores. We also include results obtained for the simple averaging, and weighted averaging when the number of used solutions corresponds to the optimal selection of the λ value.

The ensemble that achieved the best RMSE was trained on all submissions. It had non-zero coefficients for 28 solutions submitted by nine different teams. All of those coefficients were positive. The highest total impact on predictions had the submissions of *hieuwq* (0.321), followed by *Lord of the ML* (0.189), *Cyan* (0.156), and *Dymitr* (0.151). The combined impact of the remaining teams was less than 0.187.

We analyzed residuals of the resulting predictions. Figure 3 depicts their distribution across the test data. The top-left plot 3a shows the relation between ground-truth targets and predictions for test cases from the FedCSIS 2022 Challenge. It can be seen that they are aligned along the diagonal (the red line) with relatively few outlying cases. Similarly, in the top-right part of Figure 3 there is a scatter plot of residuals with regard to predictions of the ensemble. It shows that the residuals are not evenly distributed and there are a few prediction ranges with slightly larger magnitude (variance) of prediction errors. Both of those plots also show that the ensemble is nearly unbiased. This observation was confirmed – the mean difference between the ground truths and the predictions is nearly zero, i.e., -0.00027 . However, the distribution of residuals is not Gaussian. Such a hypothesis was rejected with high confidence using the Shapiro-Wilk test⁷. This can also be seen in Figures 3c and 3d. The distribution of residuals has long tails and the frequency of high error values is far from the theoretical quantile values of the Gaussian distribution. This observation suggests that the considered ensemble model could be further improved. Herein, a thorough BrightBox-based investigation of the morphology of mistakes made by each of the aforementioned 28 solutions can be helpful.

VI. CONCLUSIONS

We discussed the idea of organizing online data science competitions, with some examples taken from our own experience. We referred to KnowledgePit – our competition platform which, besides standard functionalities, includes some advanced analytical and visualization tools. We paid a special attention to the BrightBox technology which is used by KnowledgePit to approximate and diagnose solutions submitted by the competition participants. One of the methods which can be used within the resulting KnowledgePit-BrightBox environment, refers to mixing different solutions together, which leads toward more efficient ML models, as well as additional insights with regard to particular submissions.

⁶<https://cran.r-project.org/web/packages/glmnet/glmnet.pdf>

⁷https://en.wikipedia.org/wiki/Shapiro-Wilk_test

Consequently, in order to provide even more value to the future competition sponsors, we are extending KnowledgePit’s functionalities to utilize a broader range of XAI tools for the purpose of advanced post-challenge research. We also believe that such functionalities can be useful from an education perspective, whereby KnowledgePit may be considered as a platform for assessing and improving the data science competencies at universities, as well as in companies.

We also plan to continue integration of KnowledgePit with BrightBox. In particular, we are going to develop better tools for inter-competition analysis of individual platform users. Such functionality would help us to monitor the progress and skill development of participants. It could also provide value to our industrial partners who often look for potential skilled employees. In this context, KnowledgePit could be used by our partners as a tool that facilitates the recruitment of researchers for projects related to the competition topics, and for the evaluation of job candidates. Actually, QED Software is already using KnowledgePit for such evaluations.

REFERENCES

- [1] A. Janusz, D. Ślęzak, S. Stawicki, and M. Rosiak, “Knowledge Pit – A Data Challenge Platform,” in *Proceedings of the 24th International Workshop on Concurrency, Specification and Programming, Rzeszów, Poland, September 28-30, 2015*, ser. CEUR Workshop Proceedings, Z. Suraj and L. Czaja, Eds., vol. 1492. CEUR-WS.org, 2015, pp. 191–195. [Online]. Available: http://ceur-ws.org/Vol-1492/Paper_18.pdf
- [2] A. Janusz, A. Zalewska, and D. Ślęzak, “Introducing Approximation-based Model Diagnostics into KnowledgePit – A Platform for Organizing Data Mining Challenges,” in *Book of Abstracts, the 19th International Conference on Information Processing and Management of Uncertainty in Knowledge-Based Systems (IPMU 2022), Milan, Italy, July 11-15, 2022*, S. Boffa, D. Ciucci, I. Couso, J. Medina, and D. Ślęzak, Eds., 2022, pp. 19–20. [Online]. Available: <https://ipmu2022.disco.unimib.it/wp-content/uploads/sites/86/2022/07/IPMU22-book-of-abstract.pdf>
- [3] J. L. Zimmermann, “Data Competitions: Crowdsourcing with Data Science Platforms,” in *The Machine Age of Customer Insight*. Emerald Publishing Limited, 2021, pp. 183–197. [Online]. Available: <https://doi.org/10.1108/978-1-83909-694-520211017>
- [4] C. Tauchert, P. Buxmann, and J. Lambinus, “Crowdsourcing Data Science: A Qualitative Analysis of Organizations’ Usage of Kaggle Competitions,” in *Proceedings of the 53rd Hawaii International Conference on System Sciences, HICSS 2020, Maui, Hawaii, USA, January 7-10, 2020*, 2020, pp. 1–10. [Online]. Available: <https://doi.org/10.24251/HICSS.2020.029>
- [5] A. Janusz, G. Hao, D. Kałuza, T. Li, R. Wojciechowski, and D. Ślęzak, “Predicting Escalations in Customer Support: Analysis of Data Mining Challenge Results,” in *2020 IEEE International Conference on Big Data (IEEE BigData 2020), Atlanta, GA, USA, December 10-13, 2020*, X. Wu, C. Jermaine, L. Xiong, X. Hu, O. Kotevska, S. Lu, W. Xu, S. Aluru, C. Zhai, E. Al-Masri, Z. Chen, and J. Saltz, Eds., 2020, pp. 5519–5526. [Online]. Available: <https://doi.org/10.1109/BigData50022.2020.9378024>
- [6] A. Janusz, D. Kałuza, A. Chadzyńska-Krasowska, B. Konarski, J. Holland, and D. Ślęzak, “IEEE BigData 2019 Cup: Suspicious Network Event Recognition,” in *2019 IEEE International Conference on Big Data (IEEE BigData), Los Angeles, CA, USA, December 9-12, 2019*, C. K. Baru, J. Huan, L. Khan, X. Hu, R. Ak, Y. Tian, R. S. Barga, C. Zaniolo, K. Lee, and Y. F. Ye, Eds., 2019, pp. 5881–5887. [Online]. Available: <https://doi.org/10.1109/BigData47090.2019.9005668>
- [7] A. Janusz, A. Jamiołkowski, and M. Okulewicz, “Predicting the Costs of Forwarding Contracts: Analysis of Data Mining Competition Results,” in *Proceedings of the 2022 Federated Conference on Computer Science and Intelligence Systems, Sofia, Bulgaria, September 4-7, 2022*, ser. Annals of Computer Science and Information Systems, M. Ganzha, M. Paprzycki, and D. Ślęzak, Eds., vol. 30, 2022.

- [8] A. Janusz, A. Krasuski, S. Stawicki, M. Rosiak, D. Ślęzak, and H. S. Nguyen, "Key Risk Factors for Polish State Fire Service: A Data Mining Competition at Knowledge Pit," in *Proceedings of the 2014 Federated Conference on Computer Science and Information Systems, Warsaw, Poland, September 7-10, 2014*, ser. Annals of Computer Science and Information Systems, M. Ganzha, L. A. Maciaszek, and M. Paprzycki, Eds., vol. 2, 2014, pp. 345–354. [Online]. Available: <https://doi.org/10.15439/2014F507>
- [9] M. Meina, A. Janusz, K. Rykaczewski, D. Ślęzak, B. Celmer, and A. Krasuski, "Tagging Firefighter Activities at the Emergency Scene: Summary of AAI'A'15 Data Mining Competition at Knowledge Pit," in *2015 Federated Conference on Computer Science and Information Systems, FedCSIS 2015, Łódź, Poland, September 13-16, 2015*, ser. Annals of Computer Science and Information Systems, M. Ganzha, L. A. Maciaszek, and M. Paprzycki, Eds., vol. 5, 2015, pp. 367–373. [Online]. Available: <https://doi.org/10.15439/2015F426>
- [10] A. Janusz, D. Ślęzak, M. Sikora, and Ł. Wróbel, "Predicting Dangerous Seismic Events: AAI'A'16 Data Mining Challenge," in *Proceedings of the 2016 Federated Conference on Computer Science and Information Systems, FedCSIS 2016, Gdańsk, Poland, September 11-14, 2016*, ser. Annals of Computer Science and Information Systems, M. Ganzha, L. A. Maciaszek, and M. Paprzycki, Eds., vol. 8, 2016, pp. 205–211. [Online]. Available: <https://doi.org/10.15439/2016F560>
- [11] A. Janusz, T. Tajmajer, and M. Świechowski, "Helping AI to Play Hearthstone: AAI'A'17 Data Mining Challenge," in *Proceedings of the 2017 Federated Conference on Computer Science and Information Systems, FedCSIS 2017, Prague, Czech Republic, September 3-6, 2017*, ser. Annals of Computer Science and Information Systems, M. Ganzha, L. A. Maciaszek, and M. Paprzycki, Eds., vol. 11, 2017, pp. 121–125. [Online]. Available: <https://doi.org/10.15439/2017F573>
- [12] A. Janusz, Ł. Grad, and M. Grzegorowski, "Clash Royale Challenge: How to Select Training Decks for Win-rate Prediction," in *Proceedings of the 2019 Federated Conference on Computer Science and Information Systems, FedCSIS 2019, Leipzig, Germany, September 1-4, 2019*, ser. Annals of Computer Science and Information Systems, M. Ganzha, L. A. Maciaszek, and M. Paprzycki, Eds., vol. 18, 2019, pp. 3–6. [Online]. Available: <https://doi.org/10.15439/2019F365>
- [13] A. Janusz, T. Tajmajer, M. Świechowski, Ł. Grad, J. Puczniewski, and D. Ślęzak, "Toward an Intelligent HS Deck Advisor: Lessons Learned from AAI'A'18 Data Mining Competition," in *Proceedings of the 2018 Federated Conference on Computer Science and Information Systems, FedCSIS 2018, Poznań, Poland, September 9-12, 2018*, ser. Annals of Computer Science and Information Systems, M. Ganzha, L. A. Maciaszek, and M. Paprzycki, Eds., vol. 15, 2018, pp. 189–192. [Online]. Available: <https://doi.org/10.15439/2018F386>
- [14] A. Janusz, M. Przyborowski, P. Biczuk, and D. Ślęzak, "Network Device Workload Prediction: A Data Mining Challenge at Knowledge Pit," in *Proceedings of the 2020 Federated Conference on Computer Science and Information Systems, FedCSIS 2020, Sofia, Bulgaria, September 6-9, 2020*, ser. Annals of Computer Science and Information Systems, M. Ganzha, L. A. Maciaszek, and M. Paprzycki, Eds., vol. 21, 2020, pp. 77–80. [Online]. Available: <https://doi.org/10.15439/2020F159>
- [15] M. Züfle and S. Kounev, "A Framework for Time Series Preprocessing and History-based Forecasting Method Recommendation," in *Proceedings of the 2020 Federated Conference on Computer Science and Information Systems, FedCSIS 2020, Sofia, Bulgaria, September 6-9, 2020*, ser. Annals of Computer Science and Information Systems, M. Ganzha, L. A. Maciaszek, and M. Paprzycki, Eds., vol. 21, 2020, pp. 141–144. [Online]. Available: <https://doi.org/10.15439/2020F101>
- [16] P. Przybyszewski, S. Dziwiątkowski, S. Jaszczur, M. Śmiech, and M. S. Szczuka, "Use of Domain Knowledge and Feature Engineering in Helping AI to Play Hearthstone," in *Proceedings of the 2017 Federated Conference on Computer Science and Information Systems, FedCSIS 2017, Prague, Czech Republic, September 3-6, 2017*, ser. Annals of Computer Science and Information Systems, M. Ganzha, L. A. Maciaszek, and M. Paprzycki, Eds., vol. 11, 2017, pp. 143–148. [Online]. Available: <https://doi.org/10.15439/2017F567>
- [17] M. Grzegorowski, "Massively Parallel Feature Extraction Framework Application in Predicting Dangerous Seismic Events," in *Proceedings of the 2016 Federated Conference on Computer Science and Information Systems, FedCSIS 2016, Gdańsk, Poland, September 11-14, 2016*, ser. Annals of Computer Science and Information Systems, M. Ganzha, L. A. Maciaszek, and M. Paprzycki, Eds., vol. 8, 2016, pp. 225–229. [Online]. Available: <https://doi.org/10.15439/2016F90>
- [18] A. Gosiewska and P. Biecek, "Auditor: An R Package for Model-Agnostic Visual Validation and Diagnostics," *The R Journal*, vol. 11, no. 2, p. 85, 2019. [Online]. Available: <https://doi.org/10.32614/rj-2019-036>
- [19] A. Skowron and D. Ślęzak, "Rough Sets Turn 40: From Information Systems to Intelligent Systems," in *Proceedings of the 2022 Federated Conference on Computer Science and Intelligence Systems, Sofia, Bulgaria, September 4-7, 2022*, ser. Annals of Computer Science and Information Systems, M. Ganzha, M. Paprzycki, and D. Ślęzak, Eds., vol. 30, 2022.
- [20] S. Stawicki, D. Ślęzak, A. Janusz, and S. Widz, "Decision Bireducts and Decision Reducts – A Comparison," *International Journal of Approximate Reasoning*, vol. 84, pp. 75–109, 2017. [Online]. Available: <https://doi.org/10.1016/j.ijar.2017.02.007>
- [21] J. W. Grzymała-Busse, "Rule Induction," in *Data Mining and Knowledge Discovery Handbook, 2nd ed*, O. Maimon and L. Rokach, Eds. Springer, 2010, pp. 249–265. [Online]. Available: https://doi.org/10.1007/978-0-387-09823-4_13
- [22] M. Matraszek, A. Janusz, M. Świechowski, and D. Ślęzak, "Predicting Victories in Video Games – IEEE BigData 2021 Cup Report," in *2021 IEEE International Conference on Big Data (Big Data), Orlando, FL, USA, December 15-18, 2021*, Y. Chen, H. Ludwig, Y. Tu, U. M. Fayyad, X. Zhu, X. Hu, S. Byna, X. Liu, J. Zhang, S. Pan, V. Papalexakis, J. Wang, A. Cuzzocrea, and C. Ordonez, Eds., 2021, pp. 5664–5671. [Online]. Available: <https://doi.org/10.1109/BigData52589.2021.9671650>
- [23] D. Ruta, M. Liu, L. Cen, and Q. H. Vu, "Diversified Gradient Boosting Ensembles for Prediction of the Cost of Forwarding Contracts," in *Proceedings of the 2022 Federated Conference on Computer Science and Intelligence Systems, Sofia, Bulgaria, September 4-7, 2022*, ser. Annals of Computer Science and Information Systems, M. Ganzha, M. Paprzycki, and D. Ślęzak, Eds., vol. 30, 2022.
- [24] H. Xiao, Y. Liu, D. Du, and Z. Lu, "An Approach for Predicting the Costs of Forwarding Contracts using Gradient Boosting," in *Proceedings of the 2022 Federated Conference on Computer Science and Intelligence Systems, Sofia, Bulgaria, September 4-7, 2022*, ser. Annals of Computer Science and Information Systems, M. Ganzha, M. Paprzycki, and D. Ślęzak, Eds., vol. 30, 2022.
- [25] Q. H. Vu, L. Cen, D. Ruta, and M. Liu, "Key Factors to Consider when Predicting the Costs of Forwarding Contracts," in *Proceedings of the 2022 Federated Conference on Computer Science and Intelligence Systems, Sofia, Bulgaria, September 4-7, 2022*, ser. Annals of Computer Science and Information Systems, M. Ganzha, M. Paprzycki, and D. Ślęzak, Eds., vol. 30, 2022.
- [26] S. Pioroński and T. Górecki, "Using Gradient Boosting Trees to Predict the Costs of Forwarding Contracts," in *Proceedings of the 2022 Federated Conference on Computer Science and Intelligence Systems, Sofia, Bulgaria, September 4-7, 2022*, ser. Annals of Computer Science and Information Systems, M. Ganzha, M. Paprzycki, and D. Ślęzak, Eds., vol. 30, 2022.
- [27] M. Aché, A. Janusz, K. Żbikowski, D. Ślęzak, M. Kryszkiewicz, H. Rybiński, and P. Gawrysiak, "ISMIS 2017 Data Mining Competition: Trading Based on Recommendations," in *Foundations of Intelligent Systems - 23rd International Symposium, ISMIS 2017, Warsaw, Poland, June 26-29, 2017, Proceedings*, ser. Lecture Notes in Computer Science, M. Kryszkiewicz, A. Appice, D. Ślęzak, H. Rybiński, A. Skowron, and Z. W. Raś, Eds., vol. 10352. Springer, 2017, pp. 697–707. [Online]. Available: https://doi.org/10.1007/978-3-319-60438-1_68
- [28] T. Hastie, R. Tibshirani, J. H. Friedman, and J. H. Friedman, *The Elements of Statistical Learning: Data Mining, Inference, and Prediction (2nd Edition)*. Springer, 2009.