

Explaining CBR Systems Through Retrieval and Similarity Measure Visualizations: A Case Study

Paola Marín-Veites  and Kerstin Bach 

Department of Computer Science
Norwegian University of Science and Technology (NTNU)
Trondheim, Norway
`paola.m.veites@ntnu.no`

Abstract. Explainability in AI is becoming increasingly important as we delegate more safety-critical tasks to intelligent decision support systems. Case-Based Reasoning (CBR) systems are one way to build such systems. Understanding how results are created by a CBR system has become an important task in their development process. In this work, we present how visualizations can help developers and domain experts to evaluate the CBR systems behavior and provide insights to further develop CBR systems in their application scenarios. This paper presents an overview of SupportPrim, a CBR system for the management of musculoskeletal pain complaints, and presents methods that explain its retrieval and similarity measures through visualizations that help to evaluate the system's performance. In the case study, we conduct experiments within the SupportPrim CBR system using differently weighted global similarity measures to compare their effect on the retrieval. This work shows that providing suitable explanations for the CBR system's stakeholders increases the likelihood of its adoption, and visualizations allow the creation of different explanations for the different users throughout the development phase, thus allowing for better modeling and usage of the system.

Keywords: Explainable AI, XCBR, Similarity Modeling, Retrieval, Visualization

1 Introduction

Case-Based Reasoning (CBR) is an artificial intelligence method that provides a solution based on past experiences. In CBR, a case is defined as a problem description with its solution. A CBR system finds a solution by matching the new problem description (the query) to all or some of the existing cases in the case base. For searching a solution, similarity measures are used to compare problem descriptions. Once one or more solutions are found, they can be adapted or directly provided to the users. Successful cases are learned by the CBR system to increase its competence [1]. In this process the definition of the similarity measure is central as it determines which cases are returned.

Defining similarity measures and assessing them are key elements in CBR systems development. However, a CBR system usually involves users outside the technical domain that have little to no knowledge about how a CBR system is built. Stakeholders in a CBR application might understand the core concepts of CBR, but not how it works on their data. The gaps between the known methodology and how it is used in an application in their own domain can become blurry. If transparency is achieved, stakeholders and researchers would fully understand how their CBR system is operating and be able to tackle its shortcomings. This task should not only be the developer's responsibility. In many cases when creating the initial CBR system, we do not have a gold standard to compare the system against, but still need to show that the system is performing well. Visual tools can help to explore whether the retrieval differentiates well between the cases.

CBR systems are developed to serve a certain purpose, once the CBR system has been deployed, we would like to know how the CBR systems evolves over time. This includes possible updates on the similarity measures. To address this topic, we will use a CBR application example and present visualizations to create explanations on its behavior, so stakeholders understand how their data is being used within the system. For our application example, we modeled our CBR system using the local and global similarity principle.

Our overall aim is to improve the analytical tools of CBR systems and make its contents transparent to the stakeholders during the software development process. Our visualization approach can be applied to CBR systems where a concept, similarity measures, a set of attributes and a query case exist. The challenges we are addressing in this paper are as follow: (1) we visualize how the similarity of attributes contribute to an overall similarity score. (2) We present a visual explanation of the similarity scores for the query cases by comparing the global and local similarity measures of the attributes.

This paper is structured as follows: in Section 2 we discuss relevant work. Section 3 describes our application domain in more detail, in Section 4 we explain our approach and data processing, in Section 5 we test and evaluate it, in Section 6 we discuss the results and finally, our conclusions and future work are presented in Section 7.

2 Related Work

Explanations generated from CBR systems are important so the end users are encouraged to use and adopt them. As Kenny et al. mention [9], adoption barriers can be addressed by the explanation capabilities designed to improve adoption, such as adequate predictions and providing “personalised explanation-by-example”. They identify three main challenges that systems have. The first one is accuracy, the second is their interpretability for the users and the third is that they function as a good decision support system regarding its context of application. On their work, Bach et al. [4] focus on methods to explain the similarity-based retrieval reasoning process, using visualizations that allow for a

better understanding of the system, thus enabling explanations. Sørmo et al. [18] also discuss the transparency of the reasoning process and making the usage of the result understandable for the end user. They present a framework based on important explanation goals so that the context of the application aligns with it. Cunningham et al. [5] outline their experiment setting on a case-based explanation system where the decision process and patterns were transparent. In their work, subjects score the explanations. The case-based explanation system showed to perform better than having no explanation and better than rule-based systems. They concluded that the explanation potential of CBR “could have significant impact” [5]. Hoffman et al. [7] present another example on how explaining the retrieval process used in Process-Oriented Case-Based Reasoning—using graphs and workflows—helps to pinpoint how different approaches used in CBR perform. They also mention the potential for optimizing these approaches. Lamy et al. [10] propose a CBR user interface where they provide visual explanations. The interface provides both qualitative and quantitative visualizations of the similarity scores between the query case and the respective retrieval results. This approach explains why these cases were similar through visualizations that allow for easier visual reasoning. When tested with medical experts they found out that these visualizations made the shared patient characteristics easier to understand.

For the modeling and development of a CBR system, the open source case-based reasoning tool myCBR offers various explanation capabilities, as described by Roth-Berghofer et al. [16]. It is a prototyping tool with a similarity-based retrieval engine. Combining the existing tools of the software with our approach of visual explanations helps in answering questions that might arise for myCBR from the users involved in the development, such as the quality of retrieval outcome and the system’s modeling behavior. Visualizations allow to understand the system configuration as a whole, not just for the knowledge engineer. As Roth-Berghofer et al. also point out, explanations for developer engineers and explanations for the knowledge experts are different: the former is more interested in modeling errors while the latter in understanding the system’s concepts and behavior. Moreover, CBR itself has also been used as a tool to recommend explanations, for example for image classifiers [15] or to explain and assess the confidence of black-box methods [14,6].

3 SupportPrim CBR System

The CBR system used as a case study in this paper has been developed as part of the SupportPrim project, a collaborative research between the Department of Public Health and Nursing (ISM) and the Department of Computer Science (IDI) at the Norwegian University of Science and Technology (NTNU). The system’s goal is to improve management of musculoskeletal pain disorders (MSD) in primary care and provide decision support for clinical practice. In particular, we focus on creating a dashboard for clinicians that provides an overview of the

patients current situation based on previously answered questionnaires and an assessment by the clinician¹.

The SupportPrim project is a spin-off of a previous project called FysioPrim. As mentioned by Jaiswal et al. [8], the knowledge experts involved “extracted non-specific MSD patients”. The dataset is a collection of features that describe the problem for which the patient seeks consultation, such as, classification of pain areas, questionnaire responses to assess followup information and treatment, if there is a perceived improvement in the patients quality of life and functionality during treatment, as well as the patients demographics and physiological factors to mention a few.

In a second step, we use CBR to find the most similar, successful treatment plans and combine them with best practices to create an individual treatment plan. The dashboard facilitates co-decision making between clinician and patient as they are able to review the data and treatment plan together. In Fig. 1 we show how data is collected from patients and used to create datasets for our CBR systems. Questionnaires are sent out to patients before their first visit to collect baseline data. The baseline questionnaire is comprehensive and contains the items listed in Table 1 plus additional ones which are not included in the similarity matching. After 2, 4 and 8 weeks a short questionnaire including from three and up to seven questions is sent. Those questionnaires are repetitions of baseline questions and help obtaining a trajectory during treatment. After three months a larger follow-up questionnaire is deployed².

When developing the CBR system for finding most similar patients at each time point, we worked with domain experts to determine relevant attributes. Meisingset et al. [13] proposed classifying MSD patients according to phenotypes, making subgroups where treatment is adapted to the similar characteristics and prognostic factors. Attributes defining phenotypes are also relevant to find similar patient cases and we therefore used them in the case representation.

3.1 Data

For the creation of the datasets and modeling in myCBR we extracted the data collected in the SupportPrim Randomized Controlled Trial (RCT). As shown in Fig. 1 we have six datasets used to build the CBR systems: one for each time point and the demographics that can be added to extend the case representation. The datasets contains the 571 patients assessed in the different time points mentioned above. For the experiment setting, we only considered the intervention group (50% of the patients) and among those, all that have completed all stages of the intervention. Patients with missing data at each stage were dropped, which results in different case base sizes for each time point.

¹ For more information about the dashboard and the project, please see <https://www.ntnu.no/supportprim>

² The full Randomized Controlled Trial (RCT) registration can be found at <https://www.isrctn.com/ISRCTN17927832>

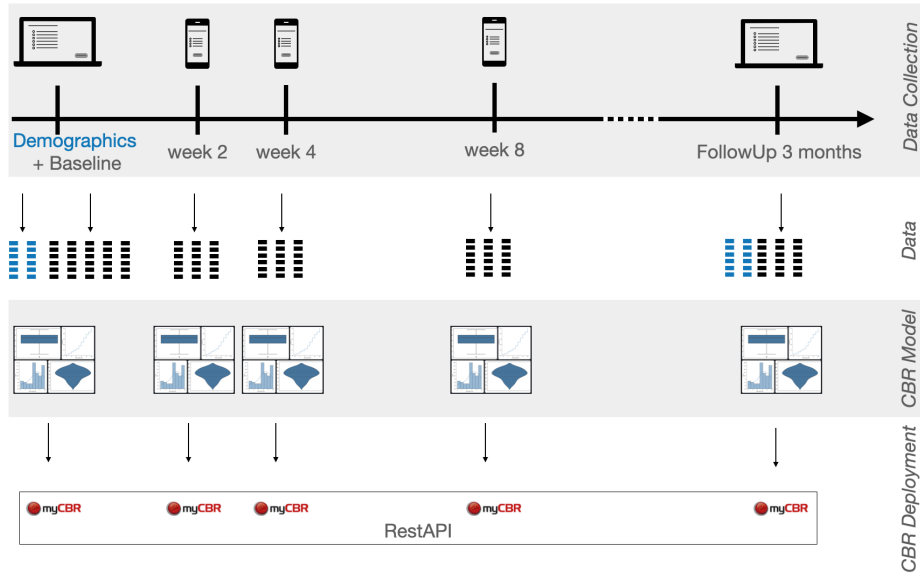


Fig. 1. Data collection and CBR system development

Table 1 contains all attributes included in the case representations. The table includes the description of the attribute, the weight distribution (i.e. its importance) used in each global similarity and their value ranges. The attributes have their origin in questionnaires used to assess patient reported outcome measures. Together with domain experts we selected relevant questionnaires for the SupportPrim study and defined relevant attributes for comparing similar patients.

Table 1: The case representation of a patient. This is a subset of the patient case dataset used in the similarity measure.

Description	Equal Weights	Different Weights	Value Range
Case ID	0.0	0.0	0 - 100000
Patient ID	0.0.	0.0	1 - ∞
Age	1.0	1.0	0 - 150
Gender	1.0	1.0	female, male
BMI	1.0	1.0	0.00 - 100.1
Smoking	1.0	1.0	no, yes
Education	1.0	2.0	primary school, high school, up to 4 years higher education, more than 4 years higher education, other
Main complaint for seeking GP	1.0	4.0	neck, shoulder, back, hip, knee, multisite

Continued on next page

Table 1 – continued from previous page

Description	Equally Weighted	Different Weighted	Value Range
Daily activity level	1.0	2.0	not reduced, slightly reduced, quite reduced, very reduced
Activity and function	1.0	1.0	0 - 10
Walk aid	1.0	4.0	no, yes
Work situation	1.0	1.0	working or other, disability pension or work assessment, sick leave
Work characteristic	1.0	2.0	mostly seated, much walking, much walking and lifting, heavy work using the body
Work ability	1.0	4.0	0 - 10
Comorbidity count	1.0	1.0	0 comorbidity, 1 comorbidity, 2 to 3 comorbidities, 4 or more comorbidities
EQ5D - Mobility	1.0	1.0	no problem, slight problem, moderate problem, severe problem, unable
EQ5D - Self-care	1.0	1.0	no problem, slight problem, moderate problem, severe problem, unable
EQ5D - Anxiety	1.0	2.0	not, slightly, moderately, severely, extremely
15D - Sleep	1.0	4.0	sleep normally, slight problem, moderate problems, great problems
15D - Vitality	1.0	1.0	healthy and energetic, slightly weary, moderately weary, very weary, extremely weary
Örebro-1: Pain Duration	1.0	4.0	less than 1 month (1-3), 1 to 3 months (4-6), 3 to 6 months (7-9), 6 to 12 months (10), more than 12 months
Örebro-Q2: Pain	1.0	1.0	0 - 10
Örebro-2: Pain last week	1.0	1.0	0 - 10
Örebro-7: Long-lasting ailments	1.0	4.0	0 - 10
Örebro-10: Stop activity	1.0	1.0	0 - 10
Number of pain sites	1.0	2.0	0 - 10
Temporary pain	1.0	1.0	no, yes
Mental distress	1.0	8.0	1.0 - 4.0
Keele STarT MSK	1.0	4.0	low, medium, high
MSK-HQ-7: Social activities and hobbies	1.0	1.0	not at all, slightly, moderately, severely, extremely
MSK-HQ-15: Physical activity level	1.0	2.0	none, 1 day, 2 days, 3 days, 4 days, 5 days, 6 days, 7 days
Pain self-efficacy Q1	1.0	1.0	0 - 6
Pain self-efficacy Q2	1.0	1.0	0 - 6

Continued on next page

Table 1 – continued from previous page

Description	Equally Weighted	Different Weighted	Value Range
Pain self-efficacy and fear avoidance score	1.0	2.0	0 - 12
Fear avoidance	1.0	1.0	0 - 10
Global Perceive Effect	1.0	1.0	very much improved, much improved, minimally improved, no change, minimally worse, much worse, very much worse

End of Table

3.2 Case Representation and Similarity Modeling

For the modeling of the attributes, we used the data from the 571 patients included in the RCT to ensure the broad value ranges are considered. The CBR systems contains 35 attributes considered as the most relevant by the domain experts involved in SupportPrim and were assessed through different time points in the RCT. Each time point reflect the overall evolution assessment of the patients and contains different number of attributes. The datasets contain symbolic, integer and float attribute types.

The local similarity measures are modeled by plotting each of the chosen attributes' Empirical Cumulative Distribution Function (ECDF). The ECDF allows to assess several characteristics of our dataset, such as the value range and distribution of the data. As mentioned by Scheidegger et al. domain models can be based on ECDF to simulate both behavior and data, because it gives a good reproduction of the observed measurements[17]. By using the ECDF plot we can model the local similarity measures using the data distributions from the underlying datasets.

Fig. 2 shows the similarity modeling for the attribute *work_ability_1* (left) and its corresponding ECDF plot (right). To align the value distribution and the similarity modeling in the value range of $[0, 1]$ where 1 is most similar, we inverted the ECDF plot. We took the proportion values in the y -axis and assumed that the distribution growth is proportional to the similarity distance. Considering the x -axis being the range values for the attribute and the y -axis being the ECDF, for example $x = 4$ and its corresponding $y = 0.3$ in the ECDF graph to the right, in the similarity measure we then model $x = 4$ as $1 - 0.3$ and thus, ending up with a similarity of 0.7 as illustrated in the graph to the left. This approach to the similarity measure modeling can be used since no more data will be further collected.

3.3 Case Base and Similarity Population

Six datasets were created for their respective case bases and concepts, each one corresponding to each time point assessment of the patient evolution. We only included the intervention group (about half of the 571 included patients) and

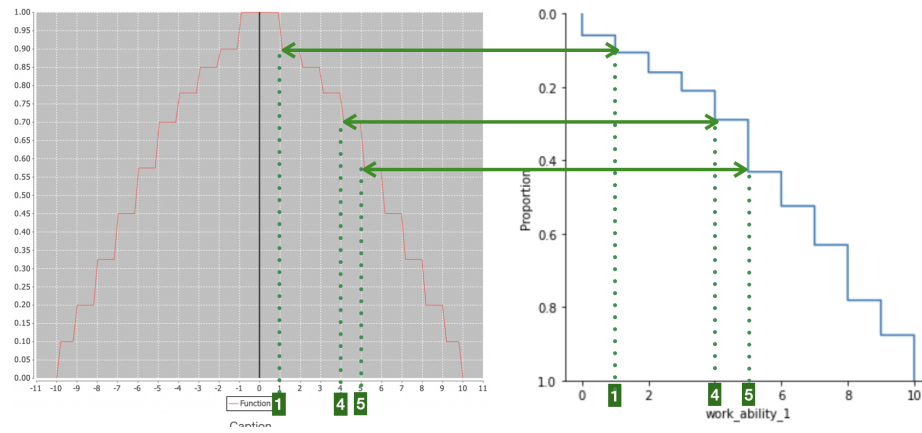


Fig. 2. Modeling of an attribute using the ECDF in the myCBR workbench

created the following case bases: demographics (269 cases), baseline (269 cases), week2 (22 cases), week4 (22 cases), week8 (22 cases), and follow-up 3 months (63 cases: see Fig. 1 for the data collection timeline).

The similarity measures were modeled in the myCBR workbench, a Java-based development framework [2,19]. It is designed to expose modeling functionality, creating concepts and similarity functions that run through an HTTP REST API and can be used with all programming languages that support Rest API and parsing JSON objects, as described by Bach et al. [3]. Local similarity measures were created for each attribute and two global similarity functions were created for each time point, one with different weights and the other equally weighted for comparison purposes. Python and Jupyter Notebook were used for the analysis and visualizations of the data facilitating the similarity modeling.

4 Explanatory Case Base Visualizations

Once the CBR agents are deployed, we can investigate the content of the case bases and how the most similar patients are retrieved. In this section we present visualizations of the CBR system that can explain a domain expert how the system is working and thereby build trust in the application. While CBR is considered to be an explainable artificial intelligence methodology [11,12], the assessment of similarity often lacks transparency. In this work, we aim to make the modeled similarity more explicit for stakeholders that are involved in the development of the CBR system. The following visualizations are not intended for end users, but for domain experts to verify that the implemented system works as intended. To create the visualizations we use the SupportPrim CBR systems including the data for the patients that have completed all stages of the intervention. This allows to compare different patient trajectories.

4.1 Accessing the CBR system’s model

For the visualizations, we created functions in Python that retrieve the data from the myCBR REST API³. The API provides access to all knowledge containers. To compare similarity functions, we used the following parameters for each time point: concept (case representation), the case base, the similarity function, a query case and k for the number of cases to retrieve in the results. We created functions to compare two instances taking into account the local similarity, the global similarity and the weights of the attributes.

4.2 Visualization of Retrievals

The retrieved results shown to the end user, e.g. the clinician and/or patient, focus on presenting cases with solutions. However, our focus is more on how the retrieval results operate on entire case bases. Fig. 3 shows the retrieval results for the SupportPrim CBR systems for all five time points. For each visualization we run a leave-one-out cross-validation using the respective case base and show the average similarity score for each ranking. Each plot contains the top 5 rank of most similar cases to the query case. The first bar to the left, $rank_0$, is the query case similarity compared to itself (to serve as a visual aid for comparison), while the following bars indicate the average similarity of the top 5 most similar cases. These five bars also show the similarity range. Such charts help to visualize how the variation of similarity can influence the results as they show the robustness of the ranking for a specific case base.

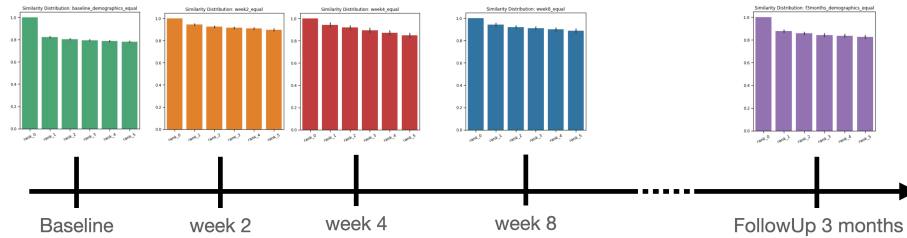


Fig. 3. Similarity score ranges for the top five cases at each time stamp

The purpose of the Fig. 3 is to show the distribution of the results per time point, providing an overview of how the system performs at each time point of the intervention. If the comparison of the mean similarity scores would be equal, we could conclude that the similarity functions are not properly representing the data distribution and the CBR system cannot differentiate relevant from

³ <https://github.com/ntnu-ai-lab/mycbr-sample-python>

irrelevant cases. In Fig. 3 we can also see the effect of different case representations to the similarity measure. The *Baseline* and *FollowUp 3 months* case representations are much larger than the weekly questionnaire case information.

4.3 Visualization of the Similarity Scores for Individual Case Comparisons

The next step after comparing the retrieval results, is to provide a visualization for understanding the detailed similarity score of a query/case pair when using the system. Fig. 4 presents a set of five case comparisons. Each comparison has three charts comparing the global similarity, the local similarity scores, and the weighted score respectively. The y -axis indicates a set of attributes, selected regarding their weights—shown in the x -axis according to Table 1—from highest (8) to lowest (1). The charts show the comparison of the five most similar cases to the query case.

The visualization shows how the weighted attributes affect the global similarity and allows to assess its influence in the overall similarity score per attribute. As mentioned by Bach et al. [4], visual explanations can offer a better understanding of the system in the development phase. With this visual aid, CBR developers can pinpoint if a set of attributes heavily weighted are overpowering other attributes' contributions.

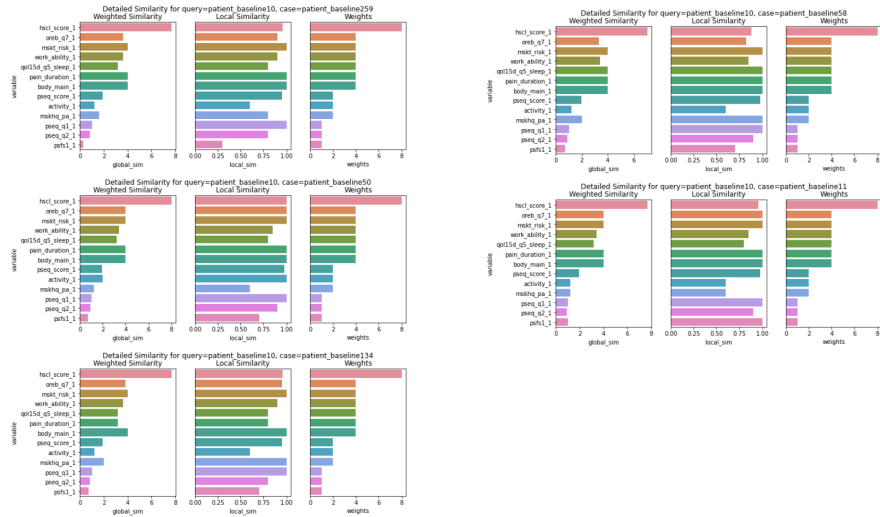


Fig. 4. Visual explanations of case comparisons. Left: rank 1 to 3; Right: rank 4 and 5

5 Experiments

In this section we will describe the experiments we conducted on the Support-Prim CBR system. When developing the local similarity measures we used the same methodology and data foundation for all time points. After discussing with our domain experts, we concluded that the assessment of patients phenotypes—as described in [13]—is a good way to compare how the CBR system performs. A phenotype describes a homogeneous patient group and we assume that they also receive similar treatment recommendations. Moreover the five phenotypes are ordered with phenotype 1 being a low risk group and phenotype 5 a high risk group. In the experiments we focus on the two largest case representation as they include all relevant attributes to compute the phenotypes, i.e. baseline and baseline plus demographics, both with $n = 269$. We conducted the experiments with both the same local similarity measures, but differently weighted global similarity measures for each case base, resulting in four different outcomes. Our case base includes five different phenotypes and the phenotypes are ordered so that neighboring phenotypes are more similar. For example, phenotypes 1 and 2 are more similar than phenotypes 1 and 3. Fig. 5 illustrates the phenotype distribution in the case base. At least 27 cases exist for each phenotype, with phenotype 5 being the smallest. Phenotype 2 is the largest with 85 cases. For the SupportPrim application, the phenotype is only assessed at baseline and hence we only conducted experiments on that CBR system.

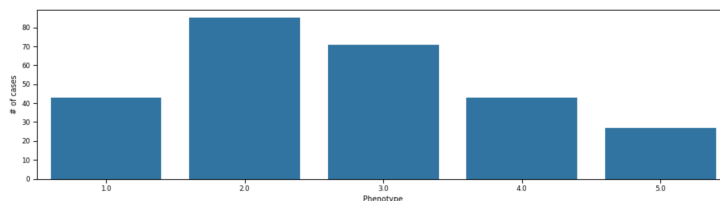


Fig. 5. Phenotype distribution in the case base

For each of the n most similar cases we calculated the difference $DiffRank_i$ between the query case’s phenotype and the case’s phenotype value. As seen from Fig. 5, we have five phenotypes so the maximum $DiffRank_i$ is 4. For each case base cb we calculated the mean phenotype error as follows:

$$MeanPhenotypeError(cb) = \frac{\sum_{i=1}^n DiffRank_i}{n} \quad (1)$$

The *MeanPhenotypeError* of both settings (equal and different weights) were identical and are shown in Table 2. The baseline column shows the mean phenotype error as calculated with Eq. 1. Additionally to the mean we also included the standard deviation.

Table 2. Results of the experiments using the baseline CBR system showing the error and standard deviation of the top n cases compared to their respective phenotypes.

	Baseline	Baseline + Demographics
Top 1 Error	0.565 (± 0.738)	0.625 (± 0.751)
Top 2 Error	0.651 (± 0.591)	0.677 (± 0.61)
Top 3 Error	0.685 (± 0.543)	0.669 (± 0.54)
Top 5 Error	0.714 (± 0.451)	0.714 (± 0.51)

The results show that using the best matching case provides on average the best results. However, the larger standard deviation indicates that there is quite some variation between the phenotypes. The variation, however, is at most of one phenotype.

6 Discussion

For the SupportPrim CBR system, we used the visualizations presented above to assess how each attribute is contributing to the overall score and to the similarity ranking among the cases. The visual aid helped in understanding the system’s performance and corroborate that the modeling of the attributes was adequate, since the matching cases are very similar to the query case. The retrieval visualizations per time point allowed to assess our case bases differentiation distributions, through the bar plots. In the application presented, we see that the cases’ similarity scores are well differentiated and the global similarity scores’ comparison charts further help assess how the different weighted global similarity is influencing the results compared to the equally weighted similarity per time point. Since the phenotypes differed by at most one, there is room for improvement. Furthermore, our evaluation results showed that the weights of the attributes in global similarity measures are not a factor in our CBR system, as the equally weighted global similarity and the different weighted global similarity had the same mean phenotype error. Visualization charts can be a tool that if paired with other analysis, e.g. correlations, can help to further explain a CBR system’s behavior. Having these pointers allows to discuss the results with domain experts and if necessary make the appropriate modifications.

7 Conclusion

From the application presented on this paper, the insights provided by the visualizations help in explaining the overall CBR system’s behavior and performance. In the development phase, having such understanding is key for modeling and assessing the similarity measures to achieve the desired CBR system goal. Visualizations allow to create explanations for different users through the development phase. The charts presented are not only information for development

but also for providing explanations to knowledge experts. A visual layout of how experts' data is used within the system can be valuable in further developing and/or updating similarity measures in CBR. Future work we would like to explore is creating visualizations on attributes' correlations within a CBR system. Although we presented how different weights of global similarity measures influenced the overall similarity score, if we visualize how the attributes are correlated with each other we can further pinpoint how each attribute, specially highly weighted ones, influences the global similarity score, and assess if a CBR system is double matching on the same feature. Moreover, creating a tool for exploring the CBR system for domain experts and developers is another way forward.

References

1. Aamodt, A., Plaza, E.: Case-Based Reasoning: Foundational Issues, Methodological Variations, and System Approaches. *Artificial Intelligence Communications*, 7(1), 39–59 (1994)
2. Bach, K., Althoff, K.D.: Developing Case-Based Reasoning Applications Using myCBR 3. In: Watson, I., Agudo, B.D. (eds.) *Case-based Reasoning in Research and Development, Proceedings of the 20th International Conference on Case-Based Reasoning (ICCBR-12)*. pp. 17–31. LNAI 6880, Springer (Sep 2012), event-place: Heidelberg, Berlin
3. Bach, K., Mathisen, B.M., Jaiswal, A.: Demonstrating the mycbr rest api. In: *Workshops Proceedings for the Twenty-seventh International Conference on Case-Based Reasoning*. vol. 2567. CEUR-WS (2016)
4. Bach, K., Mork, P.J.: On the explanation of similarity for developing and deploying CBR systems. In: *The Thirty-Third International Flairs Conference (2020)*
5. Cunningham, P., Doyle, D., Loughrey, J.: An evaluation of the usefulness of case-based explanation. In: *International conference on case-based reasoning*. pp. 122–130. Springer (2003)
6. Gates, L., Kisby, C., Leake, D.: Cbr confidence as a basis for confidence in black box systems. In: Bach, K., Marling, C. (eds.) *Case-Based Reasoning Research and Development*. pp. 95–109. Springer International Publishing, Cham (2019)
7. Hoffmann, M., Malburg, L., Klein, P., Bergmann, R.: Using siamese graph neural networks for similarity-based retrieval in process-oriented case-based reasoning. In: *International Conference on Case-Based Reasoning*. pp. 229–244. Springer (2020)
8. Jaiswal, A., Bach, K., Meisingset, I., Vasseljen, O.: Case representation and similarity modeling for non-specific musculoskeletal disorders-a case-based reasoning approach. In: *The Thirty-Second International Flairs Conference (2019)*
9. Kenny, E.M., Ruelle, E., Geoghegan, A., Shalloo, L., O'Leary, M., O'Donovan, M., Keane, M.T.: Predicting grass growth for sustainable dairy farming: A cbr system using bayesian case-exclusion and post-hoc, personalized explanation-by-example (xai). In: *International Conference on Case-Based Reasoning*. pp. 172–187. Springer (2019)
10. Lamy, J.B., Sekar, B., Guezennec, G., Bouaud, J., Séroussi, B.: Explainable artificial intelligence for breast cancer: A visual case-based reasoning approach. *Artificial intelligence in medicine* 94, 42–53 (2019)
11. Leake, D., Mcsherry, D.: Introduction to the special issue on explanation in case-based reasoning. *Artif. Intell. Rev.* 24, 103–108 (10 2005)

12. Leake, D.B.: CBR in Context: The Present and Future. In: *Case-Based Reasoning: Experiences, Lessons and Future Directions*. pp. 3–30. MIT Press (1996)
13. Meisingset, I., Vasseljen, O., Vøllestad, N.K., Robinson, H.S., Woodhouse, A., Engbretsen, K.B., Glette, M., Øverås, C.K., Nordstoga, A.L., Evensen, K.A., et al.: Novel approach towards musculoskeletal phenotypes. *European Journal of Pain* 24(5), 921–932 (2020)
14. Nugent, C., Cunningham, P.: A case-based explanation system for black-box systems. *Artif. Intell. Rev.* 24, 163–178 (10 2005)
15. Recio-García, J.A., Parejas-Llanovarcad, H., Orozco-del Castillo, M.G., Brito-Borges, E.E.: A Case-Based Approach for the Selection of Explanation Algorithms in Image Classification. In: Sánchez-Ruiz, A.A., Floyd, M.W. (eds.) *Case-Based Reasoning Research and Development*. pp. 186–200. Springer International Publishing, Cham (2021)
16. Roth-Berghofer, T.R., Bahls, D.: Explanation capabilities of the open source case-based reasoning tool mycbr. In: *Proceedings of the thirteenth UK workshop on Case-Based Reasoning UKCBR*. pp. 23–34 (2008)
17. Scheidegger, M., Baumgartner, F., Braun, T.: Simulating large-scale networks with analytical models. *International Journal of Simulation Systems, Science and Technology Special Issue on: Advances In Analytical And Stochastic Modelling* (2005)
18. Sørmo, F., Cassens, J., Aamodt, A.: Explanation in case-based reasoning—perspectives and goals. *Artificial Intelligence Review* 24(2), 109–143 (2005)
19. Stahl, A., Roth-Berghofer, T.R.: Rapid prototyping of cbr applications with the open source tool mycbr. In: *European conference on case-based reasoning*. pp. 615–629. Springer (2008)