Machine learning and case-based reasoning
TDT4173- Assignment 2

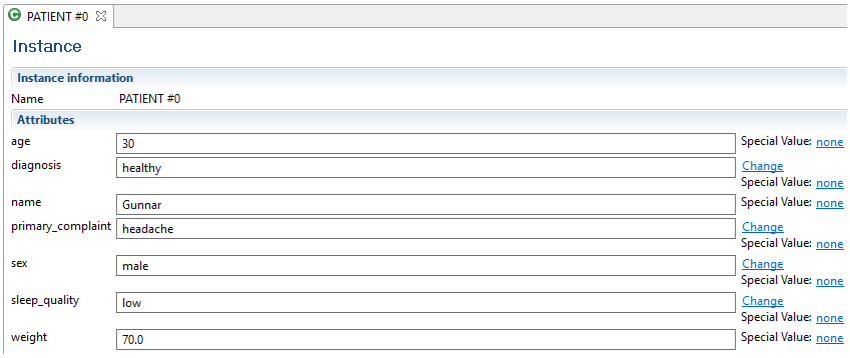
Øyvind Aukrust Rones, NTNU

February 28, 2018

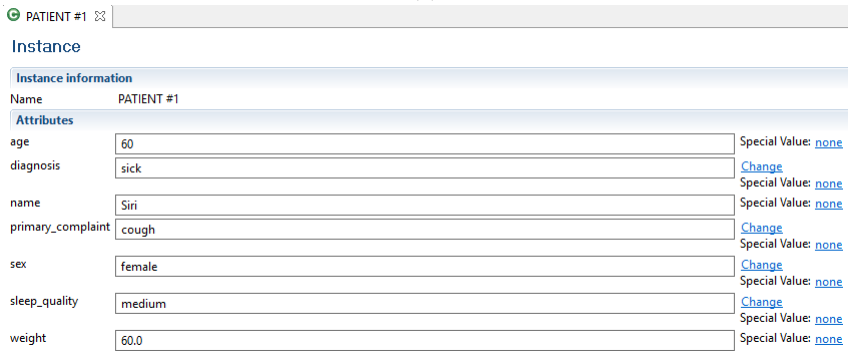
1 Practical

1. In case based reasoning, one assumes that similar problems have similar solutions. They are also *lazy* in the sense that most of the processing is done when query instances are observed. It is represented through the four step process:
   - **Retrieve** - Retrieve all cases relevant to the target problem
   - **Reuse** - Map the given solution to the target problem and adapt if needed
   - **Revise** - Revise an element of the proposed solution
   - **Retain** - Store the result as a new case in memory
2. Just like the human mind, CBR solves problem based on similar previous experiences and some assumptions. The fact that observed mistakes will result in a revision of the attempted solution is another similarity.
3. Surface similarity compares the problem with each case by looking at surface features. Structural similarities depend on the case representation and may compare objects with different structures.
4. The similarity of two cases can be considered by for example looking at the local similarities and then doing a weighted add of them.
5. Knowledge containers store different aspects of a system needed to execute the different steps of the CBR-model. Four major knowledge containers are used in CBR:
   - **Similarity Measure** - Stores the similarity between cases
   - **Case Base** - Stores all the experienced cases
   - **Adaption knowledge** - Stores the rules for adapting retrieved cases when needed
   - **Vocabulary** - Contains the attribute types and objects that the case base, adaption knowledge and similarity measures are based on.

2 Case Modeling



(a)



(b)

3 Case Retrieval

1. Figure 2 shows a ten year old male suffering from a cough. The top three matches has a similarity score between 0.82 and 0.7 although none of them has complained about coughing. This is because the local similarity between other attributes is high. For instance, all matches suffer from low sleep quality and are diagnosed as sick, while the top two matches are close in weight and age. For the primary complaint to matter more for the global similarity, a different weight could have been chosen for this attribute.

   Figure 5 shows a bit of an unexpected result. The lowest similarity of this retrieval is as high as 0.45, but this is probably because of badly weighted local similarities.

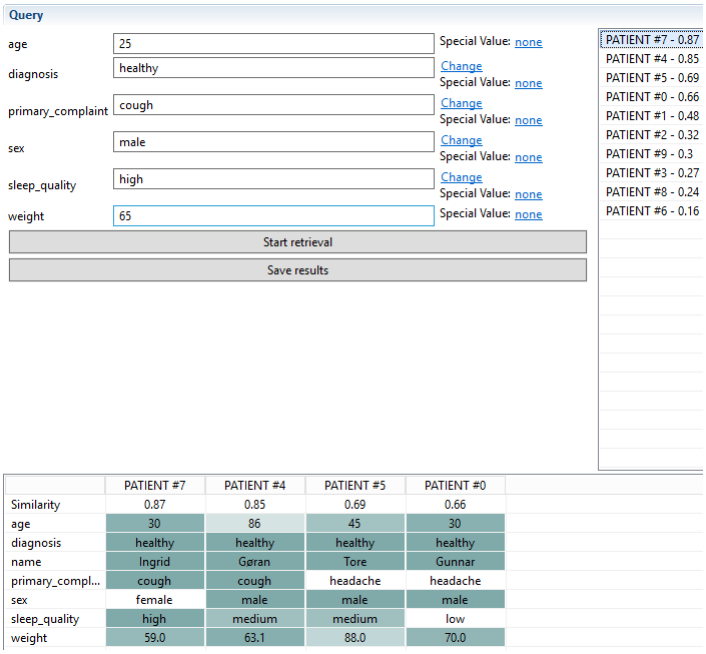
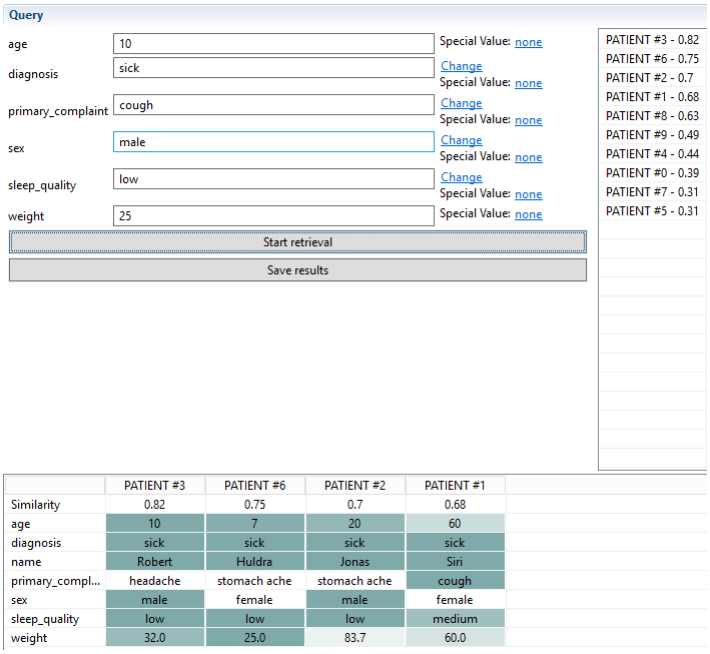


Figure 1: First retrieval

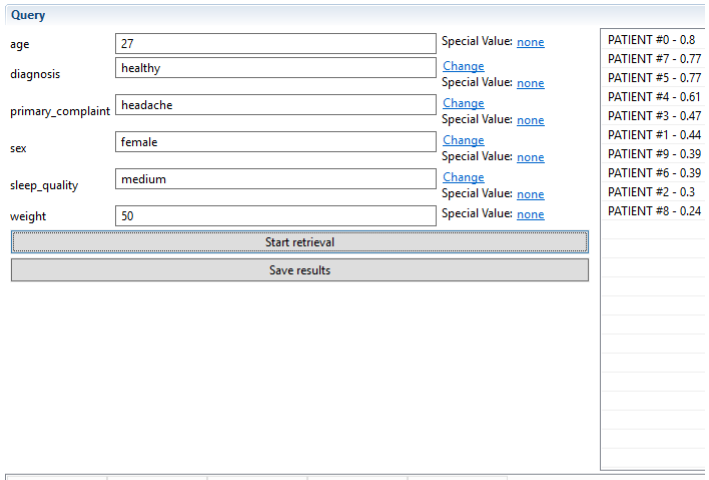
2. The PATIENT concept can be used by hospitals to automatically diagnose patients, investigate the correlation between sleep quality and weight or if a persons gender is related to health.

   The retrieval step is easily done in myCBR through the retrieval function and local-global similarity functions. The reuse step can be done by looking at all the best matches together and remove the attributes that don't add to the solution or add more attributes. In the revision step, one can for example consider if the primary complaint should be cough instead of a stomach ache. The retain step is done by adding the solution to the case base.



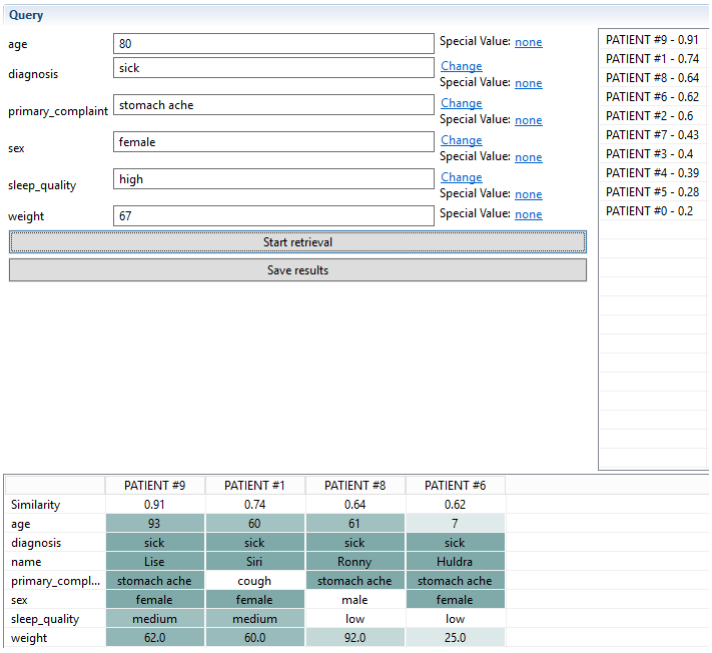
	PATIENT #3	PATIENT #6	PATIENT #2	PATIENT #1
Similarity	0.82	0.75	0.7	0.68
age	10	7	20	60
diagnosis	sick	sick	sick	sick
name	Robert	Huldra	Jonas	Siri
primary_compl...	headache	stomach ache	stomach ache	cough
sex	male	female	male	female
sleep_quality	low	low	low	medium
weight	32.0	25.0	83.7	60.0

Figure 2: Second retrieval



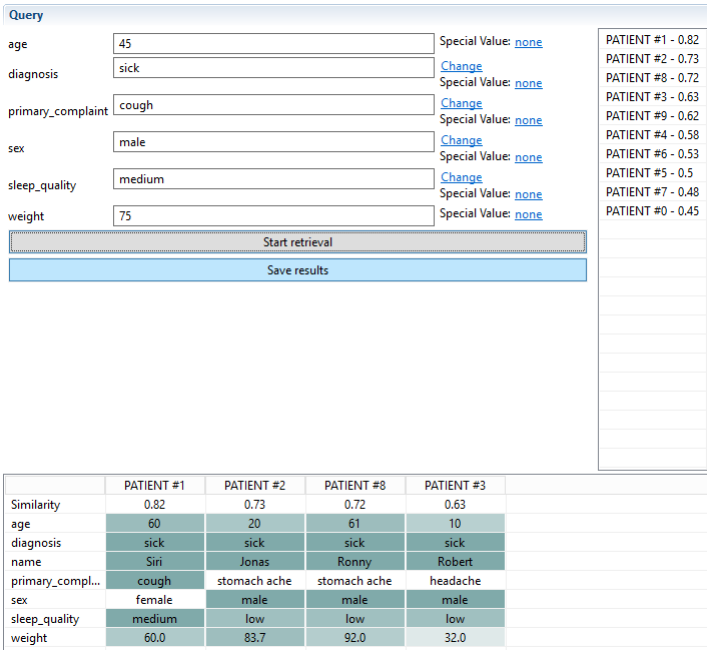
	PATIENT #0	PATIENT #7	PATIENT #5	PATIENT #4
Similarity	0.8	0.77	0.77	0.61
age	30	30	45	86
diagnosis	healthy	healthy	healthy	healthy
name	Gunnar	Ingrid	Tore	Gøran
primary_compl...	headache	cough	headache	cough
sex	male	female	male	male
sleep_quality	low	high	medium	medium
weight	70.0	59.0	88.0	63.1

Figure 3: Third retrieval



	PATIENT #9	PATIENT #1	PATIENT #8	PATIENT #6
Similarity	0.91	0.74	0.64	0.62
age	93	60	61	7
diagnosis	sick	sick	sick	sick
name	Lise	Siri	Ronny	Huldra
primary_compl...	stomach ache	cough	stomach ache	stomach ache
sex	female	female	male	female
sleep_quality	medium	medium	low	low
weight	62.0	60.0	92.0	25.0

Figure 4: Fourth retrieval



	PATIENT #1	PATIENT #2	PATIENT #8	PATIENT #3
Similarity	0.82	0.73	0.72	0.63
age	60	20	61	10
diagnosis	sick	sick	sick	sick
name	Siri	Jonas	Ronny	Robert
primary_compl...	cough	stomach ache	stomach ache	headache
sex	female	male	male	male
sleep_quality	medium	low	low	low
weight	60.0	83.7	92.0	32.0

Figure 5: Fifth retrieval