Further information on lesson 5

Notes on multi-level decision trees

When creating a multi-level decision tree, the aim is to create further rules that hierarchically follow the first rule in order to successively reduce the number of incorrect classifications. The other rules are created based on the partial data sets that were created by the first decision rule. If only green or only red labels can be found in one of the partial data sets ("pure" partial data set), no further rule needs to be searched for this partial data set. Otherwise, work continues in each branch of the tree with a different partial data set in order to set up further decision rules.

Detailed instructions on how to use the game plan when creating multi-level decision trees

The use of the game plan provides a structure for the process - the procedure can also be carried out without the game plan. The use of the game plan organizes and underlines the work in the partial data sets. Strong pupils may not need the game plan, so it can be used as a means of differentiation.

* Both partial data sets from the first data split are stored appropriately on the game plan part 1.
* Now take the first sub-group (e.. the left sub-pile of cards, NOT all the cards) and sort them on the table in ascending order according to another characteristic. The characteristic can be chosen by the pupils themselves. The unused cards remain on the game board.
* A threshold value is found and noted using the ruler method and the cards are placed in the two matching partial data sets on the game board part 2.
* Now take the second subgroup (NOT all cards, but those from level 1 that have not yet been used) and select a characteristic other than that of level 1. The same characteristic as in the left sub-data set may be used. The corresponding cards are sorted in ascending order.
* A threshold value is found using the ruler method for the second subgroup. It can be the same or a different value than in the left subgroup. The cards are placed on the game board part 2.
* Pupils can document the finished tree on AB 5.

Optional activity - Living statistics for the second level:

A typical mistake made by students is that they continue working with the complete data set in the second stage of the tree and do not use partial data sets as would be correct. Working in partial data sets can be illustrated with living statistics. However, it is up to the teacher to decide whether this "effort" is necessary for the learning group or whether a discussion in plenary is sufficient

**Living statistics**

In the first stage, the energy feature and a threshold value, e.g. 230 kcal, are used and in the second stage the protein (or sugar) feature is tested.

Procedure:

1. All pupils take a data card again.
2. The pupils divide into two groups (partial data sets) with regard to energy and the threshold value selected in the first living statistics.
3. The status quo can be discussed first
   * The number of misclassifications is counted together.
   * It is determined that the group with the higher number of calories (above the threshold value) still contains both recommended and non-recommended foods. This means that errors still occur there and the number of errors should now be reduced by a further data split.
4. Explicitly point out that only one partial data set will be used for the second data split and that the remaining cards/SuS will remain as they are (for now).
5. In the group (partial data set) with the number of calories above the threshold value, a further data split is carried out (e.. the protein characteristic with the threshold value 11 g is suitable).
6. With this second rule, all misclassifications are counted and compared with the previous number for the first level only. The second number should now be lower than the first.
7. Conclusion: The second stage has improved the tree with regard to the objective of creating a tree with as few misclassifications as possible.
8. The tree with two stages is documented on the board. The board contains a tree diagram like the one on AB5 (this finished tree can be used later, so it should be permanently preserved, e.g. on a poster).