Related datasets proposal and integration example

The integration with smart fridges is neither required in this project nor feasible since it is a fresh technology and still a niche. Nevertheless, adapting the data structure from an eventual smart fridge to the one of the ingredients table, could already be enough to store the quantity, the remaining quantity and the expire date of ingredients. The technology to store items in a smart fridge is left to the relative producer. As for the recipes part, if new ingredients need to get inserted to the system, it should be enough to adapt the data structure of those to the already existing data provided in this project.

A very important information that is not present at all in the source data is the cuisine type (French, Japanese, thai, …). A useful way to get this information is provided now. It is possible to predict the cuisine type given the ingredients. To do that, a training set is required, containing for a set of ingredients a type of cuisine. Such a training set with XX examples was found on Kaggle. Using Rapidminer, auto-modelling was used with different kind of models, like naïve bayes and decision trees. After training these models on the training set itself with a classic splitting into train and test set, the naïve bayes got a hugely higher accuracy than decision trees (around 77% against 20%). The ingredient table is the one that needed to be used to predict the cuisine type of recipes. It got adapted to the training dataset and then the prediction algorithm got started. It is not possible to have an accuracy of the results, but checking manually a couple of predictions, it seems to work quite well. This procedure is explained more precisely now.

Bla blab la blab la, model, bla bla bla, adapting, bla bla bla, predicting, bla bla bla.