A Baysian Approach to Collaborative Dish Selection

Team 10

Introduction

As anyone who has ever planned a catered event can attest, attempting to satisfy the various palates, dietary requirements and tastes of a group of diners can be a daunting task. This is particularly true given the exponential number of dishes which can be created from a small number of ingredients, as well has hard constraints such as allergies and religious beliefs. Many professional catering services handle this problem by allowing guests to select from a very limited menu. We introduce a dish recommendation system based on Bayesian Networks modeling user preferences. We predict the meals from a data base of recepices that most likely match the varied tastes of the customers, using a limited set of ingredients. This type of expert system would be of great use to a catering service or restaurant which needs to rapdily decide on a small number of dishes which would be acceptable for a large dinner party, given diverse requirements and preferences.

Related Work

Boekel and Corney propose using Bayesian Networks to model consumer needs in food production chains [5] [1]. Janzen and Xiang propose an intelligent refrigerator capable of generating meal plans based on inventory and past food choices [2]. Baysian networks have also been applied to recommendation systems before in on-line social networks [4] making predictions of the form “if you bought those items what is the probability you would like to buy that”. We suggest that these approaches are limited in that they only consider the preferences of a single (or supposed ’typical’) user rather than a group.

Approach

The approached problem is to pick a single meal which best meets the requirements and tastes of different people dining together. We learn a predictive baysian net from a survey distributed to participants of the meal as training data in order to capture their preferences. The dishes in the questionaire are selected such that all ingrediants are covered. The participants rate each dish on a scale from one to ten and give additional information like vegetarians. For new dishes we then predict the maximum likelihood rating given our model. In the following we will describe our approach in detail. First we will discuss the data selection, then the modeling of the user preference and in the end how to train the modeled net from gathered data and howe to predict the value for a new recepice.

Data accuisition We accumulated a diverse collection of sample recipes using the open source AnyMeal application. We converted to the freely available Meal-Master format (flat file) recipes to XML format for input into our application. We will gathered data representing several diners’ preference for approximately 20 meals using a simple survey of the type ’rate on a scale of 1 to 10, 10 being favorite and 1 being least favorite’. Furthermore we collected data for vegetarians and vegans.

Knowledge Engineering We will model each individual user’s preferences and needs as a Bayesian network, which means a set of independence and conditional independence relationships between variables [3]. Our model consists of 4 layers, each modeling a different aspect of taste and needs. In the first layer we capture general meal preferences, like being vegetarian or not liking your food steamed. The second layer models a general preference towards different food categories like vegetables or beef. As one can see, the food categories are dependent on the general meal preference. For example being vegetarian will exclude beef and will support vegetables. The third category models different ingredients. Each ingredient is conditioned by the food category it belongs to. In the last layer we have hard constraints like allergies (that will exclude a particular ingredient) or the overall calorie content of the meal given someone suffers from diabetes. The overall net is shown in Figure 1. Given a recipe with a list of ingredients $I = i_1, ..., i_n$ and a Bayesian network capturing user preferences we can calculate the probability of users liking the dish given the probabilities of liking each ingrediant.

Learning and Predicting In order to estimate the model parameters, the system will be trained with statistics about taste and preferences given a set of dishes with ratings from multiple users. From that information we can directly calculate the probabilities

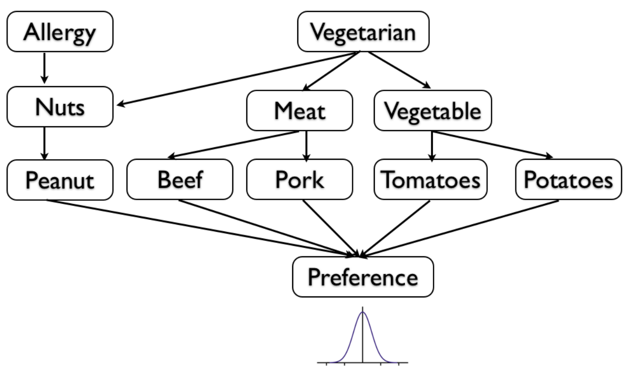

Figure 1. Our Baysian net modeling user preferences

for the ingredients using Maximum Likelihood Learning [**?**].

In order to model food preferences, we implemented a baysian net library in java. The library uses the sum-product algorithm for inference and maximum likelihood learning for parameter estimation. In our implementation we support discrete as well as continous probability distributions. Discrete distributions can be modeled as tables or as trees. In our implementation only continous distributions with discrete parents are supported. A continous distribution is then modeled as a mapping of all possible combination of it' s parents to a gaussian. Given a data set, the parameters of a discrete variable X are estimated as

$$P(X = x|Y_1 = y_1, ...Y_2 = y2) = \quad (1)$$

$$\frac{N(X = x|Y_1 = y_1, ...Y_2 = y2)}{N(Y_1 = y_1, ...Y_2 = y2)} \quad (2)$$

where $N(A)$ is the number of times event Aoccurs in the data set. We decided to implement our own Library, so we understand what is going on and we can debug and fix the models and algorithms easily.

Evaluation

The application model will be trained using a sparse subset (25-50%) of the survey data and the optimization problem soled for the inferred constraints. Next, we will calculate the correlation between the application's ranking of all dishes and the actual ranking as determined by the user surveys. We suggest that a high degree of correlation indicates that the system has the potential to accurately appraise constrained group food preferences for dishes which are not part of the survey, given sufficiently detailed recipe information.

References

[1] David Corney. Designing Food with Bayesian Belief Networks. *ACDM*, 2000.

[2] M. Janzen and Y. Xiang. Probabilistic Reasoning in Meal Planning in Intelligent Fridges. *16th Conference of the Canadian Society for Computational Studies of Intelligence*, 2003.

[3] S. Russel and P. Norvig. *Artificial Intelligence; A Modern Approach.* Prentice Hall, third edition, 2010.

[4] Tran The Truyen, Dinh Q. Phung, and Svetha Venkatesh. Preference Networks: Probabilistic Models for Recommendation Systems. *Proceedings of the Sixth Australasian Conference on Data Mining*, 2007.

[5] Stein A. van Boekel, M.A.J.S. and A.H.C. van Bruggen. Bayesian Statistics and Quality Modelling in the Agro-food Production Chain. *Proceedings of the Frontis workshop*, 2004.