Report from Dagstuhl Seminar 14101

Preference Learning

Edited by

Johannes Fürnkranz[1], Eyke Hüllermeier[2], Cynthia Rudin[3], Scott Sanner[4], and Roman Słowiński[5]

1 **TU Darmstadt, DE**
2 **Universität Marburg, DE**
3 **MIT – Cambridge, US**
4 **NICTA – Canberra, AU**
5 **Poznań University of Technology, PL**

Abstract

This report documents the program and the outcomes of Dagstuhl Seminar 14101 "Preference Learning". Preferences have recently received considerable attention in disciplines such as machine learning, knowledge discovery, information retrieval, statistics, social choice theory, multiple criteria decision making, decision under risk and uncertainty, operations research, and others. The motivation for this seminar was to showcase recent progress in these different areas with the goal of working towards a common basis of understanding, which should help to facilitate future synergies.

Seminar March 2–7, 2014 – http://www.dagstuhl.de/14101
1998 ACM Subject Classification I.2.6 Learning, J.4 Social and Behavioral Sciences
Keywords and phrases machine learning, preference learning, preference elicitation, ranking, social choice, multiple criteria decision making, decision under risk and uncertainty, user modeling, recommender systems, information retrieval
Digital Object Identifier 10.4230/DagRep.4.3.1

1 Executive Summary

Johannes Fürnkranz
Eyke Hüllermeier

The topic of "preferences" has recently attracted considerable attention in Artificial Intelligence (AI) research, notably in fields such as autonomous agents, non-monotonic reasoning, constraint satisfaction, planning, and qualitative decision theory. Preferences provide a means for specifying desires in a declarative way, which is a point of critical importance for AI. Drawing on past research on knowledge representation and reasoning, AI offers qualitative and symbolic methods for treating preferences that can reasonably complement hitherto existing approaches from other fields, such as decision theory. Needless to say, however, the acquisition of preference information is not always an easy task. Therefore, not only are modeling languages and suitable representation formalisms needed, but also methods for the automatic learning, discovery, modeling, and adaptation of preferences.

It is hence hardly surprising that methods for learning and constructing preference models from explicit or implicit preference information and feedback are among the very recent research trends in disciplines such as machine learning, knowledge discovery, information

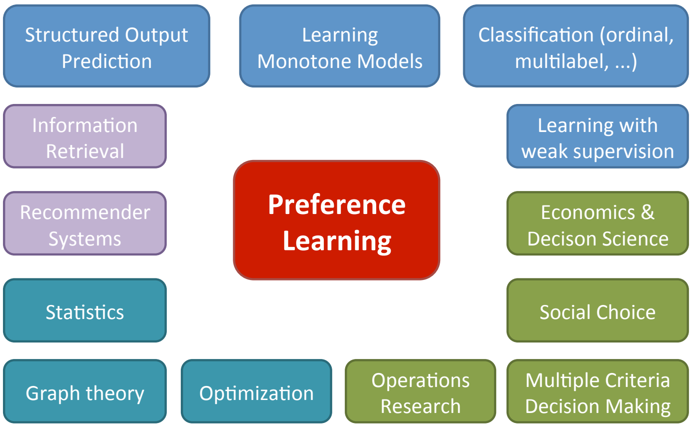

Figure 1 Preference and learning and related research areas within machine learning (blue), information retrieval (purple), applied mathematics (turquoise), and the decision sciences (green).

retrieval, statistics, social choice theory, multiple criteria decision making, decision under risk and uncertainty, operations research, and others. In all these areas, considerable progress has been made on the representation and the automated learning of preference models. The goal of this Dagstuhl Seminar was to bring together international researchers in these areas, thereby stimulating the interaction between these fields with the goal of advancing the state-of-the-art in preference learning. Topics of interest to the seminar include

- quantitative and qualitative approaches to modeling preference information;
- preference extraction, mining, and elicitation;
- methodological foundations of preference learning (learning to rank, ordered classification, active learning, learning monotone models, . . .)
- inference and reasoning about preferences;
- mathematical methods for ranking;
- applications of preference learning (web search, information retrieval, electronic commerce, games, personalization, recommender systems, . . .).

The main goal of the seminar was to advance the state-of-the-art in preference learning from a theoretical, methodological as well as application-oriented point of view. Apart from that, however, we also hope that the seminar helped to further consolidate this research field, which is still in an early stage of its development. Last but not least, our goal was to connect preference learning with closely related fields and research communities (cf. Figure 1).

In order to achieve these goals, the program featured the following components:

- Monday was filled with 6 tutorial-type introductory talks about the use of preferences and the view on preference learning in the areas of machine learning, recommender systems, multi-criteria decision making, business and economics, artificial intelligence, and social choice, with the goal of familiarizing the members of the different communities with the basics of the other fields.
- Ten sessions were devoted to contributed presentations, each one with enough extra time for discussion. In case we ran over time, we gave priority to discussions. We were also able to flexibly integrate a few impromptu talks by participants.
- Two discussion sessions on Tuesday and Thursday afternoon were devoted to discussion how to establish closer connections between the different research areas that participated in this seminar.
- Wednesday afternoon featured a hike and an excursion to Trier with some wine tasting.

2 Table of Contents

3 Introductory Talks

The program started on Monday with an entire day of introductory talks that had the goal of familiarizing the audience with each other's backgrounds.

- E. Hüllermeier, J. Fürnkranz: Preference Learning as a Machine Learning Discipline
- D. Jannach: Preference Learning in Recommender Systems – an Application-oriented Perspective
- R. Słowiński: Preference Modeling in Operational Research & Multiple Criteria Decision Aiding
- D. Baier: Preference Learning in Business and Economics: a Tutorial on Conjoint Analysis
- K. Brent-Venable, F. Rossi, T. Walsh, J. Lang: Preferences in Artificial Intelligence and Social Choice

3.1 Preference Learning as a Machine Learning Discipline

Eyke Hüllermeier (Uni Marburg) and Johannes Fürnkranz (TU Darmstadt)

The primary goal of this tutorial talk is to provide a survey of the field of preference learning in its current stage of development. Preference learning refers to the task of learning to predict an order relation on a collection of objects (alternatives). In the training phase, preference learning algorithms have access to examples for which the sought order relation is (partially) known. Depending on the formal modeling of the preference context and the alternatives to be ordered, one can distinguish between various problems types, most notably object ranking and label ranking. Both types of problems can be approached either by modeling the binary preference relation directly, or by inducing this relation indirectly via an underlying (latent) utility function.

The presentation will focus on a systematic overview of different types of preference learning problems, methods and algorithms to tackle these problems, the computational complexity of preference learning, and metrics for evaluating the performance of preference models induced from data. Along the way, we shall also try to establish a unified terminology and, moreover, to indicate connections to related research areas as well as potential applications. We will particularly focus on the aspects that are typical for machine learning, such as generalization to unseen data, and the definition of suitable loss functions which on the one hand allow to measure the learning success, and the other hand also provide the learning algorithms with criteria that can be optimized given the available training data.

References

1 Hüllermeier, E., Fürnkranz, J., Cheng, W., & Brinker, K. (2008). Label ranking by learning pairwise preferences. Artificial Intelligence, 172(16-17), 1897–1916.
2 Fürnkranz, J., & Hüllermeier, E. (Eds.). (2011). Preference Learning. Springer-Verlag.

3.2 Preference Learning in Recommender Systems – an Application-oriented Perspective

Dietmar Jannach (TU Dortmund, DE)

Main reference D. Jannach, M. Zanker, A. Felfernig, G. Friedrich, "Recommender Systems – An introduction," Cambridge University Press, ISBN 9780521493369, 2010.
URL http://www.recommenderbook.net/

The introductory talk provided an overview of common approaches to building recommender systems. Key techniques such as collaborative filtering and content-based filtering as well as knowledge-based approaches were discussed. A particular focus of the talk was on preference acquisition and learning in the context of recommender systems. The talk ended with a discussion of recent topics in the field, practical challenges, and open issues in the context of the empirical evaluation of recommender systems in research settings.

References

1 Francesco Ricci, Lior Rokach, Bracha Shapira, Paul B. Kantor (Eds.): Recommender Systems Handbook. Springer 2011
2 Jannach, D., Lerche, L., Gedikli, G., Bonnin, G.: What recommenders recommend – An analysis of accuracy, popularity, and sales diversity effects, 21st International Conference on User Modeling, Adaptation and Personalization (UMAP 2013), Rome, Italy.

3.3 Preference Modeling in Operational Research & Multiple Criteria Decision Aiding

Roman Słowiński (Poznań University of Technology, PL)

Decision problems considered in Operational Research often involve a set of alternatives (actions, objects) having vector evaluations, with the aim of either choosing the best alternative, or ranking them, or classifying them into some pre-defined and ordered classes. The vector evaluations correspond to multiple dimensions on which the alternatives are described: a dimension can be either a judgment of a voter, or an evaluation criterion, or a probability of an outcome. The three types of dimensions correspond to decision problems considered within Social Choice Theory, Multiple Criteria Decision Aiding, and Decision under Risk & Uncertainty, respectively. As evaluations on multiple dimensions are usually in conflict, the challenge consists in aggregation of evaluations on these dimensions, so as to arrive at a satisfactory recommendation formulated in terms of either the best choice, or ranking, or classification. For all these decision problems, the only objective information that stems from the problem formulation is the dominance relation in the set of alternatives. The dominance relation is, however, a partial preorder, thus it leaves many alternatives non-comparable. To enrich this relation and comparability between alternatives, a particular decision maker (DM) has to reveal her/his value system through some preference statements. This information is then used to construct/learn a preference model of the DM. This model can have the form of a synthetic value (utility) function, or a binary (outranking) relation, or a set of monotonic "*if ..., then ...*" decision rules. The preference model is inducing a preference relation on the set of alternatives. A proper exploitation of this relation leads to a recommendation [1].

We concentrate on reviewing methodologies for constructing/learning the above mentioned three types of preference models in Multiple Criteria Decision Aiding (MCDA). Moreover, we are focusing on constructing preference models from preference information provided by the DM in terms of decision examples, e.g., pairwise comparisons of some alternatives, or assignment of some alternatives to classes, or rank related requirements, or comparisons of pairs of some alternatives with respect to intensity of preference. For preference models having the form of a value function or an outranking relation, we describe a representative MCDA methodology, called Robust Ordinal Regression (ROR). ROR implements an interactive preference construction paradigm, which should be perceived as a mutual learning of the model and the DM [2, 3]. An important feature of ROR is identification of all instances of the preference model that are compatible with the input preference information – this permits to draw robust conclusions regarding DM's preferences when any of these models is applied on the considered set of alternatives. As value function models may have more or less complex form, getting a parsimonious model, adequate to the complexity of the provided preference information, is desirable.

Another aspect related to decision examples constituting the preference information is inconsistency of these examples with respect to dominance. To deal with, a Dominance-based Rough Set Approach (DRSA) has been proposed, that aims at structuring preference information into sufficiently consistent and excessively inconsistent, prior to induction of monotonic "*if ..., then ...*" decision rules considered as a logical preference model [3].

References

1 B. Roy, R. Słowiński: Questions guiding the choice of a multicriteria decision aiding method. EURO Journal on Decision Processes (2013) 1:69–97
2 J. Figueira, S. Greco, R. Słowiński: Building a set of additive value functions representing a reference preorder and intensities of preference: GRIP method. European J. Operational Research 195 (2009) 460–486.
3 S. Corrente, S. Greco, M. Kadziński, R. Słowiński: Robust ordinal regression in preference learning and ranking. Machine Learning 93(2-3): 381–422 (2013)
4 R Słowiński, S. Greco, B. Matarazzo: Rough-set-based decision support. In E.K. Burke & G. Kendall (eds.), Search Methodologies: Introductory Tutorials in Optimization and Decision Support Techniques, 2nd edition, Springer, New York, 2014, pp. 557–609.

3.4 Preference Learning in Business and Economics: a Tutorial on Conjoint Analysis

Daniel Baier (BTU Cottbus, DE)

The tutorial gives an overview on conjoint analysis, the most widely applied methodology for measuring and analyzing consumer preference in business and economics. The underlying concepts of the five steps (1) Selection of attributes and levels, (2) Design of hypothetical alternatives, (3) Collection of preferential responses, (4) Estimation of model parameters, (5) Choice prediction are discussed and illustrated by examples. A recent overview on 1.899 commercial applications of conjoint analysis are used to discuss open problems and current solutions.

4 Contributed Talks

4.1 Online Learning Over the Permutahedron: Full Information and Bandit Settings

Nir Ailon (Technion – Haifa, IL)

Joint work of Ailon, Nir; Takimoto, Eiji; Hatano, Kohei

Consider the following game: There is a fixed set V of n items. At each step an adversary chooses a score function $s_t : V \mapsto [0, 1]$, a learner outputs a ranking of V, and then s_t is revealed. The learner's loss is the sum over $v \in V$, of $s_t(v)$ times v's position (0th, 1st, 2nd, ...) in the ranking. This problem captures, for example, online systems that iteratively present ranked lists of items to users, who then respond by choosing one (or more) sought items. The loss measures the users' burden, which increases the further the sought items are from the top. It also captures a version of online rank aggregation.

We present an algorithm of expected regret $O(n\sqrt{OPT} + n^2)$, where OPT is the loss of the best (single) ranking in hindsight. This improves the previously best known algorithm of Suehiro et. al (2012) by saving a factor of $\Omega(\sqrt{\log n})$. We also reduce the per-step running time from $O(n^2)$ to $O(n \log n)$. We provide matching lower bounds.

In the bandit setting, the score functions s_t are not observed. Only the losses are observed. For this setting we present an algorithm with regret $O(n^{3/2}\sqrt{T})$ with per step running time $O(n^3)$. This trades off with a previous result of Cesa-Bianchi et al. who devise an algorithm of regret $O(n\sqrt{T \log n})$ using an algorithm that requires computing a nonnegative matrix permanent (a #P-Hard problem) at each step.

4.2 Efficient Optimization Approaches for Pairwise Ranking Losses

Antti Airola (University of Turku, FI)

Joint work of Airola, Antti; Pahikkala, Tapio; Salakoski, Tapio
Main reference A. Airola, T. Pahikkala, T. Salakoski, "Training linear ranking SVMs in linearithmic time using red-black trees," Pattern Recognition Letters. 32(9):1328–1336, 2011.
URL http://dx.doi.org/10.1016/j.patrec.2011.03.014

Straightforward approaches to minimizing pairwise ranking losses on scored data lead to quadratic costs. We demonstrate, that for the special cases of pairwise hinge loss (RankSVM) and pairwise least-squares loss (RankRLS), better scaling can be achieved by modeling the preferences only implicitly using suitable data structures.

Software implementations are available at

- http://staff.cs.utu.fi/~aatapa/software/RankSVM/(RankSVM) and
- https://github.com/aatapa/RLScore(RankRLS).

4.3 Revisiting Probabilistic Matrix Factorisation in Light of the Observed Ratings

Cédric Archambeau (Amazon CS Berlin GmbH, DE)

Joint work of Archambeau, Cédric; Balaji Lakshminarayanan; Guillaume Bouchard
Main reference B. Lakshminarayanan, G. Bouchard, C. Archambeau, "Robust Bayesian Matrix Factorisation," in Proc. of the 14th Int'l Conf. on Artificial Intelligence and Statistics (AISTAT'11), JMLR Proceedings, Vol. 15, pp. 425–433, JMLR.org, 2011.
URL http://www.jmlr.org/proceedings/papers/v15/lakshminarayanan11a/lakshminarayanan11a.pdf

We analyse the noise arising in collaborative filtering when formalised as a probabilistic matrix factorisation problem. We show empirically that modelling row- and column-specific variances is important, the noise being in general non-Gaussian and heteroscedastic. We also advocate for the use of a Student-t priors for the latent features as the standard Gaussian is included as a special case. We derive several variational inference algorithms and estimate the hyperparameters by type-II maximum likelihood. Experiments on real data show that the predictive performance is significantly improved.

4.4 Bayesian Methods for Conjoint Analysis-Based Predictions: Do We Still Need Latent Classes?

Daniel Baier (BTU Cottbus, DE)

Main reference D. Baier, "Bayesian Methods for Conjoint Analysis-Based Prediction: Do We Still Need Latent Classes?" in W. Gaul et al., (eds.), German-Japanese Interchange of Data Analysis Results, Part II; Studies in Classification, Data Analysis, and Knowledge Organization, Vol. 47, 103–113, Springer, 2014.
URL http://dx.doi.org/10.1007/978-3-319-01264-3_9

Recently, more and more Bayesian methods have been proposed for modeling heterogeneous preference structures of consumers (see, e.g.,[1, 2, 3]) Comparisons have shown that these new methods compete well with the traditional ones where latent classes are used for this purpose (see [4] for an overview on these traditional methods). This applies especially when the prediction of choices among products is the main objective (e.g. [5, 6, 7, 8] with comparative results). However, the question is still open whether this superiority still holds when the latent class approach is combined with the Bayesian one. This paper responds to this question. Bayesian methods with and without latent classes are used for modeling heterogeneous preference structures of consumers and for predicting choices among competing products. The results show a clear superiority of the combined approach over the purely Bayesian one. It seems that we still need latent classes for conjoint analysis-based predictions.

References

1 Allenby et al., J Mark Res 32:152–162, 1995, 35:384–389, 1998;
2 Baier and Polasek, Stud Classif Data Anal Knowl Organ 22:413–421, 2003;
3 Otter et al., Int J Res Mark 21(3):285–297, 2004).
4 Ramaswamy and Cohen (2007) Latent class models for conjoint analysis. In: Gustafsson A, Herrmann A, Huber (eds) Conjoint measurement – methods and applications, 4th edn. Springer, Berlin, pp 295–320.
5 Moore et al., Mark Lett 9(2):195–207, 1998;

6 Andrews et al., J Mark Res 39:479–487, 2002a; 39:87–98, 2002b;
7 Moore, Int J Res Mark 21:299–312, 2004;
8 Karniouchina et al., Eur J Oper Res 19(1):340–348, 2009

4.5 Preference-based Online Learning using Statistical Models: The Case of Mallows

Róbert Busa-Fekete (Universität Marburg, DE)

Joint work of Busa-Fekete, Róbert; Szörényi, Balázs; Hüllermeier, Eyke;

We address the problem of rank elicitation assuming that the underlying data generating process is characterized by a probability distribution on the set of all rankings (total orders) of a given set of items. Instead of asking for complete rankings, however, our learner is only allowed to query pairwise preferences. Using information of that kind, the goal of the learner is to reliably predict properties of the distribution, such as the most probable top-item, the most probable ranking, or the distribution itself. More specifically, learning is done in an online manner, and the goal is to minimize sample complexity while guaranteeing a certain level of confidence.

4.6 F-Measure Maximization for Thresholding a Ranking

Krzysztof Dembczyński (Poznań University of Technology, PL)

Joint work of Dembczyński, Krzysztof; Busa-Fekete, Róbert; Waegeman, Willem; Cheng, Weiwei; Hullermeier, Eyke

In many applications we are interested in retrieving top k elements from a ranking. There is, however, a problem how to determine k which can be given explicitly or defined through a threshold on utility values. The F-measure is commonly used to determine such a threshold in binary classification. When assuming independence of the ranked elements the F-measure satisfies the so-called probability ranking principle [4], i.e., the elements above the threshold have greater marginal probabilities of relevance than the elements below the threshold. We show how the situation changes in a general case without imposing the independence assumption [2]. We also discuss two frameworks for F-measure maximization [6]: the decision-theoretic approach and the empirical utility maximization. We also shortly address the problem of on-line maximization of the F-measure.

References

1 K. Chai. Expectation of F-measures: Tractable exact computation and some empirical observations of its properties. In *SIGIR*, 2005.
2 K. Dembczyński, W. Waegeman, W. Cheng, and E. Hüllermeier. An exact algorithm for F-measure maximization. In *NIPS*, volume 25, 2011.
3 M. Jansche. A maximum expected utility framework for binary sequence labeling. In *ACL*, pages 736–743, 2007.

4 David Lewis. Evaluating and optimizing autonomous text classification systems. In *SIGIR*, pages 246–254, 1995.
5 J. Quevedo, O. Luaces, and A. Bahamonde. Multilabel classifiers with a probabilistic thresholding strategy. *Pattern Recognition*, 45:876–883, 2012.
6 N. Ye, K. Chai, W. Lee, and H. Chieu. Optimizing F-measures: a tale of two approaches. In *ICML*, 2012.
7 M. Zhao, N. Edakunni, A. Pocock, and G. Brown. Beyond Fano's inequality: Bounds on the optimal F-score, BER, and cost-sensitive risk and their implications. *Journal of Machine Learning Research*, 14:1033–1090, 2013.

4.7 Cautious Label Ranking by Label-wise Decomposition

Sébastien Destercke (Technical University of Compiegne, FR)

Joint work of Destercke, Sébastien; Michael Poss; Marie-Helene Masson

In this talk, we present a method that aims at providing partial predictions in the setting of label ranking. We propose to do it through a label-wise decomposition scheme and to use imprecise probabilistic model to obtain the partial predictions. After a brief reminder of the imprecise probabilistic setting, we provide some details about our method and the way partial predictions can be obtained in a tractable way. In particular, we provide efficient methods to compute the Pareto-set of an assignment problem with imprecise costs described by convex sets (resulting from the imprecise probabilistic models). The method extends the recently proposed labelwise Decomposition of Cheng et al.[1] to accomodate partial predictions.

References

1 Weiwei Cheng, Sascha Henzgen, Eyke Hüllermeier. *Labelwise versus pairwise decomposition in label ranking*. Proceedings of Lernen Wissen Adaptivität 2013 (LWA13): 140-147, Otto Friedrich University Bamberg, Germany, 2013.

4.8 Exploiting Monotonicity Constraints for Arctive Learning in Ordinal Classification

Ad J. Feelders (Utrecht University, NL)

Joint work of Feelders, Ad J.; Barile, Nicola; Soons, Pieter
Main reference N. Barile, A. J. Feelders, "Active Learning with Monotonicity Constraints," in Proc. of the 2012 SIAM Int'l Conf. on Data Mining (SDM'12), pp. 756–767, 2012.
URL http://dx.doi.org/10.1137/1.9781611972825.65

In many applications of data mining it stands to reason that the response variable is increasing in the attributes. For example, the probability of acceptance for a loan increases with disposable income. Such relations between response and attribute are called monotone. If the class label of an object is given, then monotonicity may allow the labels of other objects to be inferred. For instance, knowing that applicant A is rejected, we can infer that applicants who score worse than A on all criteria should be rejected as well.

Given a collection of unlabeled attribute vectors, the question that arises is: for which vector should we request the class label from the expert, so that we can infer as many labels as possible?

We use the monotonicity constraint to augment the training sample with examples whose label can be inferred. The quality of a query strategy is measured by the predictive performance of models constructed on the resulting training sample. We consider a "monotone oracle" as well as an oracle that may produce labels that violate the monotonicity constraint.

The query strategies are evaluated on artificial data as well as publicly available real-life data sets.

4.9 A Decision-Maker Without Preferences

Andreas Geyer-Schulz (KIT – Karlsruher Institut für Technologie)

In this contribution we analyze a decision-maker without preferences. A decision-maker without preferences is a decision-maker which chooses an element of a choice set with equal probability. The problem is trivial, if the choice set is known a-priori. However, if the choice set (and its size n) is not known, we construct an (infinite) series of probability spaces and study the probability distribution of potential choice variants of k items out of n. We observe that, depending on n, rank reversals of choice variants occur, although the decision-maker acts completely rational (for small n). For large n, the order of the choice variants becomes stable, no further anomalies occur. We link this to the axiom of the violation of the independence of irrelevant alternatives in decision-theory. And in addition, we refer to research in marketing on the way consumer choices are modelled by a subsequent restriction of the choice set and the effect on branding on the human brain.

4.10 ConjointBench: Setting up and Analyzing Simple Conjoint Studies

Joachim Giesen (Universität Jena, DE)

Joint work of Giesen, Joachim; Mueller, Jens K.; Kaiser, Markus
URL http://theinf2.informatik.uni-jena.de/Software/ConjointBench.html

Conjoint analysis is a family of techniques that originated in psychology and later became popular in market research. The main objective of conjoint analysis is to measure an individual's or a population's preferences on a class of options that can be described by parameters and their levels. In choice based conjoint analysis preference data are obtained by observing test persons' choices on small subsets of the options. There are many ways to analyze choice-based conjoint analysis data. A simple but powerful approach is a reduction to a linear binary classification problem. We have implemented this reduction and use a linear support vector machine for solving the resulting classification problem. The implementation is available through the ConjointBench at our homepage at the university in Jena. The ConjointBench allows to set up simple conjoint analysis studies, to distribute a choice based questionnaire in a Doodle like manner, and to analyze the elicited data using a support vector machine.

4.11 Comparing Preference Learning with Robust Ordinal Regression and Multicriteria Customer Satisfaction Analysis

Salvatore Greco (University of Portsmouth, GB)

Multiple Criteria Decision Aiding (MCDA) offers a diversity of approaches designed for providing the decision maker (DM) with a recommendation concerning a set of alternatives (items, actions) evaluated from multiple points of view, called criteria. This presentation aims at drawing the attention of the Preference Learning (PL) community upon recent advances in a representative MCDA methodology, called Ordinal Regression, focalizing on two main issues: Robust Ordinal Regression (ROR), and measuring and analyzing customer satisfaction concerning a product through the MUSA-INT method. ROR learns by examples in order to rank a set of alternatives, thus it deals with a problem similar to that one considered by Preference Learning. ROR implements, however, an interactive preference construction paradigm, which should be perceived as mutual learning of the preference model and the DM, and not as discovering of a preference model preexisting in the DM's mind. The talk clarifies the specific interpretation of the concept of preference learning adopted in ROR and MCDA, and shows similarities and differences with respect to the usual concept of preference learning considered within PL. This comparison concerns the structure of the considered problem, the types of admitted preference information, the form of the employed preference models, the ways of exploiting them, and, finally, the techniques applied to arrive at a final ranking. MUSA-INT methodology generalizes the MUSA (MUlticriteria Satisfaction Analysis) method. MUSA is a preference disaggregation method that, following the principle of ordinal regression analysis, finds an additive utility function representing both the comprehensive satisfaction level of a set of customers and a marginal satisfaction level with respect to each criterion. Differently from MUSA, MUSA-INT takes also into account positive and negative interactions among criteria, similarly to the multicriteria method UTAGMS-INT. MUSA-INT accepts evaluations on criteria with different ordinal scales which do not need to be transformed into a unique cardinal scale prior to the analysis. Moreover, instead of a single utility function, MUSA-INT can also take into account a set of utility functions representing customers' satisfaction, adopting the robust ordinal regression methodology. An illustrative example shows how the proposed methodology can be applied on a customers survey.

4.12 Multidimensional Unfolding and Clustering of Preferences: A New Simulation Design

Willem J. Heiser (Leiden University, NL)

Unfolding models are built on the concept of single-peaked preference functions that have different locations on a scale or in a space of options. The key idea is to construct a joint scale or a joint space that contains two kinds of points: one set of points for the options, and another set of points for the judges, where the latter are called ideal points because they represent the position of the peak in the single-peaked preference functions, and hence the ideal option that a judge could imagine. The objective of multidimensional unfolding then is to locate the ideal points and the option points in the joint space, in such a way that their inter-point Euclidean distances are inversely related to the preferences. We discuss a particular unfolding method and program called PREFSCAL, based on least squares and optimal data transformation. Next, we present a clustering method for preferences, called Cluster Component Analysis (CCA), which is based on the Kemeny distance between rankings, and show how it can be combined with the unfolding representation. We also outline a new simulation design for generating clusters of rankings from central rankings that satisfy an unfolding model. In this type of design, we can keep the dispersion within clusters and the amount of overlap between clusters under control, while also generating noise rankings which do not satisfy the unfolding model. Our first results indicate that CCA can recover the original central rankings very well, and that the unfolding representation is also recoverable.

References

1 Willem J. Heiser and Antonio D'Ambrosio. Clustering and Prediction of Rankings Within a Kemeny Distance Framework. In B. Lausen et al. (eds.), Algorithms from and for Nature and Life, Studies in Classification, Data Analysis, and Knowledge Organization, DOI: 10.1007/978-3-319-00035-0, Springer 2013
2 Frank M.T.A. Busing, Willem J. Heiser, Graham Cleaver. Restricted unfolding: Preference analysis with optimal transformations of preferences and attributes. Food Quality and Preference 21:82–92, 2010.

4.13 Preferences in an Open World: Perspectives for Preference Learning

Ulrich Junker (Biot, DE)

Main reference U. Junker, "Preferences in an Open World," in Proc. of the 1st Int'l Conf. on Algorithmic Decision Theory (ADT'09), LNCS, Vol. 5783, pp. 215–224, Springer, 2009.
URL http://dx.doi.org/10.1007/978-3-642-04428-1_19

Decision making may involve multiple viewpoints which are comparing the given options according to different preference relations. Examples are the viewpoints of multiple agents in group decision making or the viewpoints imposed by different criteria in multi-criteria decision making. The talk studies questions that arise when multiple viewpoints are merged into a single viewpoint over a combinatorial criteria space. The talk revisits a preference model presented at the ADT 2009 conference and explores its possibilities for preference learning.

The merging of viewpoints requires an aggregation of the preferences of the individual viewpoints, for example by adopting a ceteris-paribus semantics. Preferences can thus be aggregated in a purely deductive way without requiring any additional learning step. According to this method, it is sufficient to learn the preferences of the individual agents in order to predict the decisions of a group of agents.

However, the strict ceteris-paribus semantics may turn out to be too restrictive. What happens if agents accurately follow their individual preferences in individual situations, but the decision made by a group of agents contradicts the predictions made by the preference aggregation? Such a scenario permits the learning of a new preference over the merged viewpoint that states that the observed decision is strictly preferred to the decision predicted under the ceteris-paribus semantics. This new preference will conflict with the ceteris-paribus preferences.

We present an approach that aggregates preference relations while applying the ceteris-paribus principle as a default rule instead of a strict rule. More specific preference statements over the merged viewpoints can thus override preferences resulting from aggregating the preferences of the individual viewpoints. The resulting preference model provides the same predictions as the standard model if no observation contradicts these predictions, but is able to accommodate to situations where the observations contradict the predicted behaviour. It thus provides new perspectives for preference aggregation and preference learning in combinatorial domains.

4.14 Rank Loss Minimization with Pointwise Surrogates

Wojciech Kotłowski (Poznań University of Technology, PL)

We consider the problem of rank loss minimization or, equivalently, maximization of AUC, in bipartite ranking and multilabel classification. Since the complexity of these problems is quadratic in the number of training examples/labels, it is tempting to ask how much can be done by minimizing a simple pointwise (univariate) loss function, as done by standard classification methods, as a surrogate. We show that weighted (cost-sensitive) versions of standard margin-based surrogates, such as exponential or logistic loss, are consistent for rank loss minimization. Instead of directly proving convergence, we give a stronger result by deriving regret bounds and convergence rates. The proposed losses suggest efficient and scalable algorithms, which are tested experimentally. We also extend our results to the case of rank loss minimization in multipartite ranking (ordinal regression).

4.15 Graded Multilabel Classification by Pairwise Comparisons

Eneldo Loza Mencía (TU Darmstadt, DE)

Joint work of Brinker, Christian; Loza Mencía, Eneldo; Fürnkranz, Johannes

The task in multilabel classification is to predict for a given set of labels whether each individual label should be attached to an instance or not. Graded multilabel classification generalizes this setting by allowing to specify for each label a degree of membership on an ordinal scale. This setting can be frequently found in practice, for example when movies or books are assessed on a one-to-five star rating in multiple categories.

In this paper, we propose to reformulate the problem in terms of preferences between the labels and their scales, which then be tackled by learning from pairwise comparisons. We present three different approaches which make use of this decomposition and show on three datasets that we are able to outperform baseline approaches.

In particular, we show that our solution, which is able to model pairwise preferences across multiple scales, outperforms a straight-forward approach which considers the problem as a set of independent ordinal regression tasks.

4.16 A Brief Survey on Learning Compact Representations of Preferences over a Combinatorial Domain

Jérôme Mengin (Paul Sabatier University – Toulouse, FR)

We consider orderings over a combinatorial domain, for instance a catalog where items are defined by a number of options, so that the number of available items is exponential in the number of options. Can we learn an ordering of the items from observations of users navigating in this catalog, in order to guide future users of the catalog ? We survey a few results on learning two types of compact representations for this ordering.

Generalized additive utilities rank the items according to the sum of their scores on a limited number of subsets of the options. Such a representation is easy to learn from examples of pairwise comparisons when the structure (the subsets of options) are known, but learning the structure is hard.

Conditional preference rules of the form "if X is the case, then this value for option Y is preferred to that value" can also be used to compactly represent preferences. Reasoning with such rules can be tractable if the rules are associated with some structure over the set of options. For instance, if there is an importance, possibly partial, ordering over the set of variables, then pairwise comparisons can be done in linear time, and learning the rules can also be done in polynomial time from observations of such pairwise comparisons. CP-nets, in which is a directed graph, usually acyclic, over the set of variables represent preferential dependencies, enable fast retrieval of optimal (undominated) items, and can be learnt efficiently from observations of optimal items.

4.17 Learning Ordinal Sorting Models from Large Learning Sets: A Multicriteria Decision Aid Perspective

Vincent Mousseau (Ecole Centrale Paris, FR)

Joint work of Mousseau, Vincent; Pirlot, Marc; Sobrie, Olivier
Main reference O. Sobrie, V. Mousseau, M. Pirlot, "Learning a Majority Rule Model from Large Sets of Assignment Examples," in Proc. of the 3rd Int'l Conf. on Algorithmic Decision Theory (ADT'13), LNCS, Vol. 8176, pp. 336–350, Springer, 2013.
URL http://dx.doi.org/10.1007/978-3-642-41575-3_26

Multiple criteria sorting methods assign alternatives to predefined ordered categories. The Majority Rule Sorting model (MR-Sort) is an outranking based sorting method corresponding to a simplified version of Electre Tri. Learning the parameters of a MR-Sort model through linear programming requires the use of binary variables. In the context of preference learning where large sets of alternatives and numerous attributes are involved, such an approach is not an option in view of the large computing times implied. Therefore, we propose a new metaheuristic designed to learn the parameters of an MR-Sort model. This algorithm works in two phases that are iterated. The first one consists in solving a linear program determining the weights and the majority threshold, assuming a given set of profiles. The second phase runs a metaheuristic which determines profiles for a fixed set of weights and a majority threshold. The presentation focuses on the metaheuristic and reports the results of numerical tests, providing insights on the algorithm behavior. The perspective of handling large datasets to learn preference models is discused in the context of Multicriteria Decision Aiding.

References

1 Olivier Sobrie, Vincent Mousseau, and Marc Pirlot. Learning a majority rule model from large sets of assignment examples. In Patrice Perny, Marc Pirlot, and Alexis Tsoukiàs, editors, *Algorithmic Decision Theory – Third International Conference, ADT 2013*, volume 8176 of *Lecture Notes in Computer Science*, pages 336–350. Springer, 2013. ISBN 978-3-642-41574-6.

4.18 Making Decisions with High-Level Preferences and User-Centric Principles

Ingrid Oliveira de Nunes (Federal University of Rio Grande do Sul, BR)

Joint work of Nunes, Ingrid Oliveira de; Luck, Michael; Miles, Simon; Barbosa, Simone; Lucena, Carlos
Main reference I. O. de Nunes, "User-centric Preference-based Decision Making," PhD Thesis, Pontifícia Universidade Católica do Rio de Janerio (PUC-Rio), 2012.
URL http://www.inf.ufrgs.br/~ingridnunes/publications/0912914_2012_Completa.pdf

Choosing from a set of available options often requires resolution of trade-offs but it can be unfeasible for humans to carefully evaluate each option of a large set due to the required time and cognitive effort. Consequently, they are often unsatisfied with their choices. Software systems can support human decision making or even automate this process, but there are many challenges associated with the provision of such support. In this talk, I will first introduce a new preference meta-model founded on a study of how humans express preferences, allowing the representation of high-level preferences. Then, I will introduce an automated decision making technique, which chooses an option from a set available based on preferences expressed in a language based on the meta-model, exploiting natural-language terms. This technique makes decisions with the incorporation of psychology principles, which concern how humans make decisions, as the provided preferences are typically not enough to resolve trade-offs among available options. Finally, I will present an explanation generation technique, which uses models built by the decision making technique to justify choices, and follows guidelines and patterns derived from a study of choice explanation.

References

1 Ingrid Nunes. *User-centric Preference-based Decision Making*. PhD thesis, PUC-Rio, Brazil, 2012.
2 Ingrid Nunes, Simon Miles, Michael Luck, and Carlos Lucena. User-centric preference-based decision making (extended abstract). In Padgham Conitzer, Winikoff and van der Hoek, editors, *Eleventh International Conference on Autonomous Agents and Multiagent Systems (AAMAS'12)*, pages 1225–1226, Valencia, Spain, June 2012. IFAAMAS.
3 Ingrid Nunes, Simon Miles, Michael Luck, and Carlos Lucena. User-centric principles in automated decision making. In Finger M. Pozo A.T.R. Gimenenez-Lugo G.A. Barros, L.N. and M. Castilho, editors, *21st Brazilian Symposium on Artificial Intelligence (SBIA 2012)*, volume 7589 of *LNCS*, pages 42–51, Curitiba, Brazil, October 2012. Springer-Verlag.
4 Ingrid Nunes, Simone Barbosa, Don Cowan, Simon Miles, Michael Luck, and Carlos Lucena. Natural language-based representation of user preferences. *Interacting with Computers*, 2013. Advanced Access.

4.19 Algorithmics of Tensor-Based Preference Learning

Tapio Pahikkala (University of Turku, FI)

Main reference T. Pahikkala, "Fast Gradient Computation for Learning with Tensor Product Kernels and Sparse Training Labels," in Proc. of the IAPR Int'l Workshop on Structural and Syntactic Pattern Recognition (S+SSPR 2014), to appear; preprint available from the author's webpage.
URL http://staff.cs.utu.fi/~aatapa/publications/Pahikkala2014SSSPR.pdf

We consider the problem of learning utility functions and rankings with paired inputs and tensor-based kernel functions defined on such inputs. With paired inputs, we refer to the ones consisting of a condition and an object part. The condition being, for example, a query object given at prediction time, the learned model assigns scores for a set of target objects also given at prediction time, that indicate the conditional utility of the targets for the query. We present a new learning algorithm for the considered setting whose computational efficiency is improved with tensor-algebraic optimization.

References

1 T. Pahikkala. Fast gradient computation for learning with tensor product kernels and sparse training labels. In *IAPR International Workshop on Structural and Syntactic Pattern Recognition (S+SSPR 2014)*, 2014. To appear.
2 T. Pahikkala, W. Waegeman, A. Airola, T. Salakoski, and B. De Baets. Conditional ranking on relational data. In *Machine Learning and Knowledge Discovery in Databases (ECML PKDD 2010)*, volume 6322 of *LNCS*, pages 499–514. Springer, 2010. 10.1007/978-3-642-15883-4_32.
3 T. Pahikkala, A. Airola, M. Stock, B. D. Baets, and W. Waegeman. Efficient regularized least-squares algorithms for conditional ranking on relational data. *Machine Learning*, 93 (2-3):321–356, 2013. 10.1007/s10994-013-5354-7.
4 T. Pahikkala, A. Airola, S. Pietilä, S. Shakyawar, A. Szwajda, J. Tang, and T. Aittokallio. Toward more realistic drug-target interaction predictions. *Briefings in Bioinformatics*, 2014. 10.1093/bib/bbu010. In press.
5 W. Waegeman, T. Pahikkala, A. Airola, T. Salakoski, M. Stock, and B. De Baets. A kernel-based framework for learning graded relations from data. *IEEE Transactions on Fuzzy Systems*, 20(6):1090–1101, December 2012. 10.1109/TFUZZ.2012.2194151.

4.20 A Borda Count for Collective Sentiment Analysis

Francesca Rossi (University of Padova, IT)

Joint work of Grandi, Umberto; Loreggia, Andrea; Rossi, Francesca; Saraswat, Vijay;
Main reference U. Grandi, A. Loreggia, F. Rossi, V. Saraswat, "From Sentiment Analysis to Preference Aggregation," in Proc. of the Int'l Symp. on Artificial Intelligence and Mathematics (ISAIM'14), 2014.
URL http://www.cs.uic.edu/pub/Isaim2014/WebPreferences/ISAIM2014_CSC_Grandi_etal.pdf

Sentiment analysis assigns a positive, negative or neutral polarity to an item or entity, extracting and aggregating individual opinions from their textual expressions by means of natural language processing tools. In this paper we observe that current sentiment analysis techniques are satisfactory in case there is a single entity under consideration, but can lead to inaccurate or wrong results when dealing with a set of possibly correlated items. We argue in favor of importing techniques from voting theory and preference aggregation to provide more accurate definitions of the collective sentiment for a set of multiple items. We propose a notion of Borda count which combines individuals' sentiment and preference information, we show that this class of rules satisfies a number of properties which have a natural interpretation in the sentiment analysis domain, and we evaluate its behavior when faced with highly incomplete domains.

4.21 Exact Bayesian Pairwise Preference Learning and Inference in Expressive Models

Scott Sanner (NICTA – Canberra, AU)

Joint work of Sanner, Scott; Abbasnejad, Ehsan
Main reference S. Sanner, E. Abbasnejad, "Exact Bayesian Pairwise Preference Learning and Inference in Expressive Models," NIPS Workshop on Choice Models and Preference Learning, 2011.
URL http://users.cecs.anu.edu.au/~sguo/cmpl2011_submission_14.pdf

In Bayesian approaches to utility learning from preferences, the objective is to infer a posterior belief distribution over an agent's utility function based on previously observed agent preferences. From this, one can then estimate quantities such as the expected utility of a decision or the probability of an unobserved preference, which can then be used to make or suggest future decisions on behalf of the agent. However, there remains an open question as to how one can represent beliefs over agent utilities, perform Bayesian updating based on observed agent pairwise preferences, and make inferences with this posterior distribution in an exact, closed-form. In this paper, we build on Bayesian pairwise preference learning models under the assumptions of linearly additive multi-attribute utility functions and a bounded uniform utility prior. These assumptions lead to a posterior form that is a uniform distribution over a convex polytope for which we then demonstrate how to perform exact, closed-form inference w.r.t. this posterior, i.e., without resorting to sampling or other approximation methods.

References

1 Scott Sanner and Ehsan Abbasnejad. *Exact Bayesian Pairwise Preference Learning and Inference in Expressive Models*. NIPS Workshop on Choice Models and Preference Learning, Sierra Nevada, Spain, 2011.

4.22 Preferences, Invariances, Optimization

Michèle Sebag (University of Paris South XI, FR)

Joint work of Loshchilov, Ilya; Schoenauer, Marc; Sebag, Michèle;

Some optimization settings deal with the user in the loop (a.k.a. interactive optimization) or with expensive ill-posed optimization objectives (e.g. in numerical engineering where the optimization objective is computed using Finite Element methods).

In such settings, the number of optimization queries should be minimized, and one way to do so is to learn an approximation of the optimization objective, referred to as surrogate model.

Note that replacing the optimization objective F with g(F), with g any monotonous function, does not harm the optimization goal. Accordingly, the surrogate model of F can be learned using preference learning.

The talk will describe how the tight integration of preference learning and the distribution-based optimization algorithm CMA-ES achieves a black-box optimization algorithm which is invariant under monotonous transformations of the optimization objective, and affine transformations of the feature space.

References

1 Arnold, L., Auger, A., Hansen, N., and Ollivier, Y. Information-Geometric Optimization Algorithms: A Unifying Picture via Invariance Principles. ArXiv e-prints, 2011.

2 Hansen, N., Ostermeier, A. Completely derandomized self- adaptation in evolution strategies. Evolutionary computation 9 (2), 159-195, 2001. Surrogate models for optimization

3 Loshchilov, I., Schoenauer, M., Sebag, M. Self-adaptive surrogate-assisted covariance matrix adaptation evolution strategy. GECCO 2012, ACM Press: 321-328, 2012.

4 Loshchilov, I., Schoenauer, M., Sebag, M. A mono surrogate for multiobjective optimization. GECCO 2010, ACM Press: 471-478.

5 Viappiani, P., Boutilier, C. Optimal Bayesian Recommendation Sets and Myopically Optimal Choice Query Sets. NIPS 2010: 2352-2360

6 Hoos, H. H. Programming by optimization. Commun. ACM 55(2): 70-80.

4.23 Multiresolution Analysis of Incomplete Rankings

Eric Sibony (Télécom Paris Tech, FR)

Joint work of Sibony, Eric; Clémençon Stéphan; Jakubowicz, Jérémie
Main reference S. Clémençon, J. Jakubowicz, E. Sibony, "Multiresolution Analysis of Incomplete Rankings," arXiv:1403.1994v1 [math.ST] , 2014.
URL http://arxiv.org/abs/1403.1994v1

Incomplete rankings on a set of items $1, \ldots, n$ are orderings of the form $a_1 < \ldots < a_k$, with $a_1, \ldots, a_k \subset 1, \ldots, n$ and $k < n$. Though they arise in many modern applications, only a few methods have been introduced to manipulate them, most of them consisting in representing any incomplete ranking by the set of all its possible linear extensions on $1, \ldots, n$. In this talk, I will introduce a completely novel approach, which allows to treat incomplete rankings directly, representing them as injective words over $1, \ldots, n$. Unexpectedly, operations on incomplete rankings have very simple equivalents in this setting and the topological structure of the complex of injective words can be interpretated in a simple fashion from the perspective of ranking. We exploit this connection here and use recent results from algebraic topology to construct a multiresolution analysis and develop a wavelet framework for incomplete rankings. Though purely combinatorial, this construction relies on the same ideas underlying multiresolution analysis on a Euclidean space, and permits to localize the information related to rankings on each subset of items. It can be viewed as a crucial step toward nonlinear approximation of distributions of incomplete rankings and paves the way for many statistical applications, including preference data analysis and the design of recommender systems.

4.24 What is a Decision Problem?

Alexis Tsoukiàs (University Paris-Dauphine, FR)

Joint work of Colorni, Alberto; Tsoukiàs, Alexis;
Main reference A. Colorni, A. Tsoukiàs, "What is a decision problem? Preliminary Statements," in Proc. of the 3rd Int'l Conf. on Algorithmic Decision Theory (ADT'13), LNCS, Vol. 8176, pp. 139–153, Springer, 2013.
URL http://dx.doi.org/10.1007/978-3-642-41575-3_11

The presentation introduces a general framework about what is a decision problem. The aim is to provide a theory under which the existing methods and algorithms can be characterised, designed, chosen or justified. The framework shows that 5 features are necessary and sufficient in order to completely describe the whole set of existing methods. It also explains why optimisation remains the general approach under which decision problems are algorithmically considered.

4.25 The Limitations of Convex Surrogate Losses for Learning to Rank

Nicolas Usunier (Technical University of Compiegne, FR)

Part of the research on learning to rank has been driven by its application to search engines, where the training data consists of user queries, candidate documents for each query, and where information on the desired ordering of the documents is obtained from user feedback or paid annotators. In that context, the community has put a great emphasis on designing algorithms that optimize a convex objective function on the training data. The exact form of the convex objective function vary from one algorithm to another, but in all cases the convex objective is used as a computationally tractable surrogate of a pre-specified quality measure of the predicted rankings. The use of convex surrogate approaches is usual in machine learning, and theoretically well- grounded for classification tasks in the sense that optimizing a well-chosen convex objective function asymptotically leads to an optimal classifier. However, as I will show in this talk, such desirable properties of convex surrogate approaches do not extend to ranking: for some of the most common quality measures used to evaluate search engines, it is impossible to generate an optimal ranking function by optimizing a convex objective function. The result implies in particular that many existing algorithms for learning to rank cannot optimize the quality measure they are designed for, even in a favorable asymptotic regime.

4.26 Incremental Elicitation of Choquet Integrals using Minimax Regret

Paolo Viappiani (UPMC – Paris, FR)

Joint work of Benabbou, Nawal; Patrice, Perny; Viappiani, Paolo;

The Choquet integral is one of the most sophisticated and expressive preference models used in decision theory for multicriteria decision making. It performs a weighted aggregation of criterion values using a capacity function assigning a weight to any coalition of criteria, thus enabling positive and/or negative interactions among criteria and covering an important range of possible decision behaviors. However, the specification of the capacity involves many parameters which raises challenging questions, both in terms of elicitation burden and guarantee on the quality of the final recommendation. In this paper, we investigate the incremental elicitation of the capacity through a sequence of preference queries selected one-by-one using a minimax regret strategy so as to progressively reduce the set of possible capacities until a decision can be made. We propose a new approach designed to efficiently compute minimax regret for the Choquet model. Numerical experiments are provided to demonstrate the practical efficiency of our approach.

4.27 User Modeling with Sparse, Implicit Feedback, e-Shop Data

Peter Vojtáš (Charles University – Prague, CZ)

Joint work of Vojtáš, Peter; Peska, Ladislav; Eckhardt, Alan; Horváth, Tomas
Main reference L. Peska, A. Eckhardt, P. Vojtáš, "UPComp – A PHP Component for Recommendation Based on User Behaviour," in Proc. of the 2011 IEEE/WIC/ACM Int'l Conf. on Web Intelligence and Intelligent Agent Technology (WI-IAT'11), Volume 3, pp. 306–309, 2011.
URL http://dx.doi.org/10.1109/WI-IAT.2011.180

In this report we extend the abstract of our Dagstuhl presentation. The extension consists of related bibliographic references (ordered by time at the end of this report) and short comments on development of our views in the field of preference learning (starting here).

Our previous research was based in fuzzy logic programming, uncertain reasoning and databases. Main impulse came from an anonymous referee at a computer science conference which asked "Where from do your rules (of fuzzy logic programs) come from?" This was an important question also because in this time I have moved to Prague to the Department of Software Engineering and we wanted to contribute to the field (at least from a broader perspective).

Our first reaction was starting research in fuzzy (many valued) inductive logic programming. When looking for data to learn from we used school rating data and were able to find dependencies between ratings of subjects. Immediately, it was clear that our fuzzy values have a comparative meaning, e.g. if physics is at least B or better then Math is at least B or better (in data we learned from). Real life (software engineering relevant) data came from understanding fuzzy degrees as degrees of preferences (inducing ordering). Most challenging were problems with multiple users and recommendation. After a period of research of learning preferences form explicit rating of users, we came to our last point of interest: learning preferences from implicit behavior of a user (typically on an e-shop).

So now, I can discuss with my software engineering colleagues problems of real applications (which classical UML modeling neglected).

Original Dagstuhl abstract. Our motivation considers recommendation in SME e-shops in area where there is a large competition. In such case users usually do not register and do not rate items. Only information we have are behavioral data collected by PHP scripts. Only direct indicator of preference is purchase. Our model is based on Fagin-Lotem-Naor [1] representation of single user preferences on attributes and aggregating them. Our task is to learn parameters for a many users variant of the FLN model (we have a many users variant of FLN top-k threshold algorithm). But we assumed there is no explicit rating and the only direct preference indicator is purchase! Because of sparseness of data, we take all purchases from all users together (collaborative aspect) and learn some generalization of dependences between their behavior and purchases. We select a t-conorm from a parameterized family and obtain a single rating (user independent) of all behaviors. A new user (test set) behavior is interpreted as explicit ratings of items visited by that user and we learn parameters of FLN model. For each user separately we get a global rating (and hence a ranking) for all items. We evaluate our method on real production data from a travel agency. Finally we report on our other projects, related works and discuss various dimensions of the decision making problem/process.

References

1. R. Fagin, A. Lotem, M. Naor: Optimal aggregation algorithms for middleware. J. Comput. Syst. Sci. 66(4): 614-656 (2003)
2. L. Peska, P. Vojtáš. Recommending for Disloyal Customers with Low Consumption Rate. SOFSEM 2014: 455–465
3. L. Peska, P. Vojtáš. Enhancing Recommender System with Linked Open Data. FQAS 2013: 483-494
4. L. Peska, P. Vojtáš. Negative implicit feedback in e-commerce recommender systems. WIMS 2013: 45
5. L. Peska, P. Vojtáš. Estimating importance of implicit factors in e-commerce recommender systems. WIMS 2012: 62
6. A. Eckhardt, P. Vojtáš. User Preference Learning in Real Systems: from Events to Processes. EJC 2011: 305-312
7. L. Peska, A. Eckhardt, P. Vojtáš. UPComp – A PHP Component for Recommendation Based on User Behaviour. Web Intelligence/IAT Workshops 2011: 306-309
8. B. Vaclav, A. Eckhardt, P. Vojtáš. Pref Shop A Web Shop with User Preference Search Capabilities. Web Intelligence/IAT Workshops 2010: 330-333
9. A. Eckhardt, P. Vojtáš. Learning User Preferences for 2CP-Regression for a Recommender System. SOFSEM 2010: 346-357
10. A. Eckhardt, P. Vojtáš. How to Learn Fuzzy User Preferences with Variable Objectives. IFSA/EUSFLAT Conf. 2009: 938-943
11. A. Eckhardt, P. Vojtáš. Evaluating Natural User Preferences for Selective Retrieval. Web Intelligence/IAT Workshops 2009: 104-107
12. A. Eckhardt, P. Vojtáš. Combining Various Methods of Automated User Decision and Preferences Modelling. MDAI 2009: 172-181
13. P. Vojtáš. Decathlon, Conflicting Objectives and User Preference Querying. DATESO 2008
14. P. Vojtáš, A. Eckhardt: Considering Data-Mining Techniques in User Preference Learning. Web Intelligence/IAT Workshops 2008: 33-36
15. A. Eckhardt, T. Horváth, Dusan Maruscak, Robert Novotny, P. Vojtáš: Uncertainty Issues and Algorithms in Automating Process Connecting Web and User. URSW (LNCS Vol.) 2008: 207-223

16 A. Eckhardt, J. Pokorný, P. Vojtáš: Integrating user and group preferences for top-k search from distributed web resources. DEXA Workshops 2007: 317-322
17 A. Eckhardt, J. Pokorný, P. Vojtáš: A System Recommending Top-k Objects for Multiple Users Preferences. FUZZ-IEEE 2007: 1-6
18 A. Eckhardt, T. Horváth, P. Vojtáš: Learning Different User Profile Annotated Rules for Fuzzy Preference Top-k Querying. SUM 2007: 116-130
19 A. Eckhardt, T. Horváth, P. Vojtáš: PHASES: A User Profile Learning Approach for Web Search. Web Intelligence 2007: 780-783
20 T. Horváth, P. Vojtáš: Induction of Fuzzy and Annotated Logic Programs. ILP 2006: 260-274
21 T. Horváth, P. Vojtáš: Ordinal Classification with Monotonicity Constraints. Industrial Conference on Data Mining 2006: 217-225
22 P. Gurský, T. Horváth, R. Novotny, V. Vanekova, P. Vojtáš: UPRE: User Preference Based Search System. Web Intelligence 2006: 841-844
23 T. Horváth, P. Vojtáš: Fuzzy Induction via Generalized Annotated Programs. Fuzzy Days 2004: 419-433
24 T. Horváth, F. Sudzina, P. Vojtáš: Mining Rules from Monotone Classification Measuring Impact of Information Systems on Business Competitiveness. BASYS 2004: 451-458.

4.28 The PeerRank Method

Toby Walsh (NICTA – Kensington, AU)

Main reference T. Walsh, "The PeerRank Method for Peer Assessment," in Proc. of the 21st Europ. Conf. on Artificial Intelligence (ECAI'14), to appear; pre-print available as arXiv:1405.7192v1 [cs.AI].
URL http://arxiv.org/abs/1405.7192v1

I propose the PeerRank method for peer assessment. This constructs a grade for an agent based on the grades proposed by the agents evaluating the agent. Since the grade of an agent is a measure of their ability to grade correctly, the PeerRank method weights grades by the grades of the grading agent. The PeerRank method also provides an incentive for agents to grade correctly. It rewards agents who grade well, and penalises those that grade poorly. As the grades of an agent depend on the grades of the grading agents, and as these grades themselves depend on the grades of other agents, I define the PeerRank method by a fixed point equation similar to the PageRank method for ranking web-pages. I identify some formal properties of the PeerRank method, discuss some examples, compare with related work and evaluate the performance on some synthetic data.

References

1 Toby Walsh. *The PeerRank Method.* Proc. of ECAI-2014, IOS Press, Prague, 2014

5 Discussions

The discussion sessions revolved around the use of preferences in various disciplines. As a result of these discussions, we have been able to establish a comprehensive survey of the properties that characterize machine learning, multi-criteria decision aid and conjoint analysis as different approaches to preference learning, showing where these fields share commonalities but also where they differ with respect to underlying assumptions, goals, and methods (cf. Table 1). This discussion helped the participants to broaden their view, and to show more plainly in which way the fields can complement and mutually benefit from each other.

As a concrete follow-up project, we decided to organize a joint special issue in the *European Journal of Operational Research* (EJOR). In order to establish a joint focus, the plan is to use an industrial dataset as a common basis for potential contributions. Thus, the idea is to collect contributions that tackle and exploit the data in different ways, employing the tools of the respective communities.

Table 1 Comparison of properties of the disciplines preference learning (PL), multi-criteria decision aiding (MCDA), and conjoint analysis (CA).

	PL	MCDA	CA
Problem focus	predictions	user/decision maker	model
User interaction	typically not, yet possible in active learning	constructive, feedback with user in the loop	prior to data collection
Learning domain	population (generalize across individuals)	single user	population
Representation of alternatives	feature-based, but also structured, often many (generic) features	monotone, well-engineered criteria, decision space versus criteria space	conjoint structure, well-engineered features
Representation of users	feature-based	no features of the DM used	feature-based
Preference information	global/holistic, example-based	local and/or global, rich specifications	local and/or global, highlighting heterogeneity
Nature of the data	noisy/probabilistic	consistent, possibly corrected	noisy/probabilistic but well designed
Models and model assumptions	possibly weak assumptions(compensated by massive data)	stronger assumptions, axiomatic foundation	interpretable, often (generalized) linear models
Model interpretation, usage, and expectations	mainly predictive, accurate prediction of decision maker's behavior	mainly constructive or normative, convincing explanations of decisions	mainly descriptive, useful descriptions of decision makers
Data availability	data sets massively available (but not always accessible)	limited, user-generated data, no benchmark data	data abounds, many practical projects
Data volume	possibly very large ("big data")	typically small	moderate
Validation, success criteria	accuracy metrics, internal validation on data	user satisfaction (difficult to measure)	external evaluation (business oriented)
Computational aspects	scalability is critical	less critical (but short response time required)	less critical
Application domains	broad but typically not safety-critical (e-commerce, etc.), automated decisions	broad, possibly safety-critical, one-shot decisions	business and marketing

Participants

- Nir Ailon
Technion – Haifa, IL
- Fabio Aiolli
University of Padova, IT
- Antti Airola
University of Turku, FI
- Cédric Archambeau
Amazon CS Berlin GmbH, DE
- Daniel Baier
BTU Cottbus, DE
- Jerzy Błaszczyński
Poznań Univ. of Technology, PL
- Róbert Busa-Fekete
Universität Marburg, DE
- Weiwei Cheng
Universität Marburg, DE
- Yann Chevaleyre
University of Paris North, FR
- Krzysztof Dembczyński
Poznań Univ. of Technology, PL
- Sébastien Destercke
Technical University of Compiegne, FR
- Ad J. Feelders
Utrecht University, NL
- Johannes Fürnkranz
TU Darmstadt, DE
- Andreas Geyer-Schulz
KIT – Karlsruhe Institute of Technology, DE
- Joachim Giesen
Universität Jena, DE
- Salvatore Greco
University of Portsmouth, GB
- Willem J. Heiser
Leiden University, NL
- Eyke Hüllermeier
Universität Marburg, DE
- Dietmar Jannach
TU Dortmund, DE
- Ulrich Junker
Biot, FR
- Kristian Kersting
TU Dortmund, DE
- Wojciech Kotłowski
Poznań Univ. of Technology, PL
- Jérôme Lang
University Paris-Dauphine, FR
- Eneldo Loza Mencá
TU Darmstadt, DE
- Jérôme Mengin
Paul Sabatier University – Toulouse, FR
- Vincent Mousseau
Ecole Centrale Paris, FR
- Ingrid Oliveira de Nunes
Federal University of Rio Grande do Sul, BR
- Alena Otto
Universität Siegen, DE
- Tapio Pahikkala
University of Turku, FI
- Marc Pirlot
University of Mons, BE
- Michael Rademaker
Ghent University, BE
- Francesca Rossi
University of Padova, IT
- Scott Sanner
NICTA – Canberra, AU
- Michele Sebag
University of Paris South XI, FR
- Eric Sibony
Télécom Paris Tech, FR
- Roman Słowiński
Poznań Univ. of Technology, PL
- Alexis Tsoukiàs
University Paris-Dauphine, FR
- Nicolas Usunier
Technical University of Compiegne, FR
- K. Brent Venable
Tulane University, US
- Paolo Viappiani
UPMC – Paris, FR
- Peter Vojtáš
Charles University – Prague, CZ
- Toby Walsh
NICTA – Kensington, AU
- Paul Weng
UPMC – Paris, FR
- Christian Wirth
TU Darmstadt, DE