Research Statement:
Machine Learning with Humans in the Loop

Karthik Raman
http://www.cs.cornell.edu/~karthik/

Intelligent systems, ranging from internet search engines and online retailers to personal robots and MOOCs, live in a symbiotic relationship with their users - or at least they should. On the one hand, users greatly benefit from the services provided by these systems. On the other hand, these systems can greatly benefit from the world knowledge that users communicate through their interactions with the system. These interactions – queries, clicks, votes, purchases, answers, demonstrations, etc. – are one of the key sources of "Big Data", providing enormous potential for economically and autonomously optimizing these systems and for gaining unprecedented amounts of world knowledge required to solve some of the hardest AI problems.

The key challenge in working with data that results from human behavior is that it typically does not fit standard machine learning models. In particular, interactions typically do not provide us with ground truth labels, but are simply the result of decisions made by humans. The core insight behind my research is that we need new machine learning models and algorithms that explicitly account for how humans make decisions, which is influenced by a variety of factors like decision context, expertise, skills, motivations and needs. Armed with this insight, I have developed new learning models and learning algorithms that provide provable performance guarantees for learning from human interaction data. I have also implemented and fielded human-interactive learning systems, such as the text search engine for `arxiv.org`, that autonomously, robustly, and cost-effectively improve their performance.

In this vast research area of learning with humans in the loop, I have so far focused on problems motivated by search, recommendation, and education analytics. This work, and how it opens up new research directions, is discussed in the following. As will be evident from this work, my goal is to study machine learning problems with real-world impact from the combined lenses of machine learning, algorithmic theory, and human interaction. I strive to come up with algorithms, that have theoretical guarantees, which I then implement on real user-facing systems and evaluate via user studies.

Going forward technological changes, such as smart homes, are only going to grow and enrich the modes for humans to interact with intelligent systems. By exploiting such advances, I hope to work on and create new disruptive technologies that can fundamentally alter our day-to-day life.

1 Revealed Preferences: Learning from User Interactions

Traditional machine learning algorithms for information systems have relied on expert annotated data (e.g., assessors are paid to rate search results on a Likert scale). However a more economical source of data is the *implicit feedback* that users provide through their interactions (e.g., clicks). The advantages of using such feedback data are clear: this feedback is not only available in abundance, but also directly indicates the users' – not the experts' – preferences. Consider, for example web search, where such feedback is readily available as users scan the results page and click on different results, which provides information about the *goal-directed* choices users make based on their preferences.

To learn from this weak feedback, we have developed *Coactive Learning* algorithms [1, 2] that explicitly incorporate models of how boundedly rational users make decisions. Coactive learning is an online model of interaction between a learning system and human user, where the goal is to maximum user satisfaction. At each step, the system (e.g. search engine) receives a context (e.g. query) from the user. The system then predicts an object (e.g. ranking) and presents it to the user. In response, the user's interaction with the system (e.g. via clicks) results in feedback about the presented object. This feedback, however, typically does not reveal what would have been the optimal object to present, but only an incremental improvement to the presented object. For example, clicks on the search results B and D for the ranking $[A, B, C, D, ...]$, can help us infer that the user would have preferred the ranking $[B, D, A, C, ...]$, but not that it is the best possible ranking. More generally and in contrast to standard machine learning where optimal feedback is required, coactive learning merely requires feedback that slightly improves on the presented object. Furthermore, this user feedback may contain noise and be biased by factors such as the presentation order.

For this weaker form of preference feedback, which is readily available from user interactions in many settings, we have developed new learning algorithms in conjunction with plausible models of user interaction. The interplay between the learning algorithm and boundedly rational users leads to learning systems that can learn on-the-fly and have strong theoretical guarantees. In particular, our algorithms converge towards the optimal solution at a rate that is proportional to the square root of the number of learning steps and independent of the dimensionality of the feature space. Note that despite being provided with noisy, biased preferences as feedback, the convergence rates achieved are asymptotically equivalent to those in the full-information case *i.e.,* when the optimal object is provided as feedback.

In addition to proving theoretical guarantees, I also established that coactive learning performs robustly and accurately in a real-world setting with real users. In particular, I built an online system that implemented these algorithms on the experimental text search engine at `arxiv.org`, a scientific repository for e-prints. Empirical studies [2] with real users demonstrated that these algorithms can learn successfully from the interactions with users, optimizing retrieval performance quickly and operating completely autonomously without any maintenance from me for over a year (and running).

Beyond coactive learning, I have also developed *Correctable Learning*, an alternate paradigm for incorporating humans in the learning loop [3]. In correctable learning, the system receives feedback about which examples it has incorrectly learned so that it can look to rectify these mistakes. We provided an efficient correctable learning algorithm that incorporates such feedback to iteratively improve the learned model by utilizing local sub-models for different regions of the space. The resulting algorithm is simpler to debug and extend than conventional learning algorithms used today.

2 Complex Utilities: Learning Diversified Recommendations

Conventional search and recommendation algorithms model the relevance of items to be independent of other items in the ranking. Clearly, this simplifying assumption is not true in many cases. One such case is diversified retrieval, where the goal is to provide a comprehensive set of results that are distinct and cover the different needs of users. Previous approaches to this problem have had to make hand-coded choices between two factors: partially satisfying all the different needs *vs.* specifically catering to the most common need. In my research, I have studied methods that learn the right balance between these two extremes by explicitly modeling the joint utility of a set of items using submodular functions.

Intrinsic Diversity (ID) is one such problem where we jointly model the relevance of a set of items. The goal of intrinsic diversification is to cover different aspects of the information need of a *single* user. For example, a user of a personalized news system would not like to see exclusively articles about the Middle East Crisis on any given day, even if this was the topic he was most interested in. Instead, a diversified portfolio of topics that covers all interests of the user maximizes the user's overall utility. To approach this problem, we developed algorithms [1] that continuously learn both the relevance of items and the appropriate amount of diversity a user desires. These algorithms learn from set-valued preference data derived from the implicit feedback of user interactions. Theoretical and empirical analysis of these algorithms again reveal convergence rates equivalent to optimal feedback conditions. These algorithms have been deployed within a recommendation system for scientific articles at `http://my.arxiv.org/arxiv`.

Intrinsic Diversity is particularly prominent in web search, since studies have shown that a large fraction of real-world search tasks are intrinsically diverse. However, since current research on web search has focused solely on optimizing and evaluating single queries, these complex tasks currently require significant user effort via multiple interactions with the search engines. An ideal search engine would not only retrieve relevant results for a user's particular query, but also be able to identify when the user is engaged in a more complex task and aid the user in completing that task. Our work was the first to study the impact of ID in web search and make progress towards the goal of optimizing whole-task relevance [4]. It was awarded the **Best Student Paper** at SIGIR 2013. In particular, we addressed three key problems for ID retrieval [5]: identifying authentic instances of ID tasks from post-hoc analysis of behavioral signals in search logs; learning to identify queries that mark the start of an ID search task; and given an ID query, improving the search experience by predicting which content to prefetch and rank using a joint model of the document relevances and aspect relevance to the underlying task

Another problem which requires the joint modeling of a set of items is *Extrinsic Diversity*. This problem arises when *different* users express different information needs via the same query, thus resulting in ambiguous queries. Diversification is now used to provide relevant results for all information needs and all users, avoiding that the most popular intent drowns out all other intents. In order to satisfy all users, we used the idea of joint relevance and diversity modeling to design coactive learning algorithms that do not require explicit feedback, but can learn from user interactions [6]. While these algorithms use similar submodular models as in the case of intrinsic diversity, we are now optimizing the ranking not just for a single utility function, but for a distribution of utility functions. In addition to strong theoretical bounds, these algorithms also display significantly faster convergence than existing algorithms for single-query diversification. Furthermore, these were the first known algorithms for the task of cross-query diversification from implicit feedback.

While extrinsically diverse rankings mitigate the problem of completely missing the user's intent, they are necessarily a compromise between the breadth and the depth of coverage of each user intent. To overcome the constraints of this compromise, we developed a new dynamic retrieval model that is not restricted to a single ranking [7]. In particular, our model replaces the single one-size-fits-all ranking with a two-level ranking, where the second order rankings are conditioned on the user's interactions with the top-level ranking Constructing these two-level rankings requires modeling diversity (in the top-level ranking) and relevance (in the second-level rankings), for which we developed a Structural SVM learning algorithm.

3 Stated Preferences: Scaling up Student Evaluation

Modeling how humans make decisions is essential not only for understanding the preferences they reveal through their actions, but also for understanding the preferences they state explicitly. As an example, I have investigated the problem of peer grading, where students grade each other's work. Peer grading is a promising approach for tackling the problem of student evaluation at scale, since the number of graders scales with the number of students. However, students are not trained graders, which motivates grading models that are more robust than asking students to assign letter grades. To this effect, I proposed eliciting *ordinal* feedback (where the emphasis is on ordering alternatives) from students as it is easier to provide and more reliable than cardinal feedback [8]. Under this feedback model, student graders make ordinal statements (e.g. *project X is better than project Y*) as opposed to cardinal statements (e.g. *project X is a B-*). We designed algorithms to aggregate this ordinal feedback from individual graders to infer an overall grade for each assignment. The proposed algorithms not only model the quality of the assignments, but also the reliability of the different graders, since graders may have differing skills and grading expertise. To demonstrate the applicability of these methods, I conducted a user study in a real class. By collecting peer grades from the students of the class I was able to demonstrate that the proposed techniques are a viable alternate to traditional evaluation techniques (instructor/TA grading). By surveying the students we were able to establish that students found the overall peer-grading process to be a helpful and valuable experience.

To further increase adoption of these techniques, I also explored Bayesian alternates to this grade-aggregation problem [9]. By computing the Bayesian posterior of the ranking distribution, course instructors receive more information about the uncertainty of each assignment's aggregated grade. This information can be used for better interpreting the resulting grades or for assigning additional graders to assignments with high posterior entropy. The resulting techniques have been deployed for use at `peergrading.org`.

4 Other Work

As mentioned in the introduction, human interactions with online systems are one of the key sources of Big Data. To make processing such data tractable, including the many preprocessing augmentation steps that are required for analysis, it is common practice to build pipelines of data processing tasks. However, these big-data pipelines are currently constructed in a rather ad-hoc manner. To remedy this, I have explored modeling data-processing pipelines as graphical models, so as to improve inference [10] and enable more robust and efficient goal-oriented learning.

In addition to extracting knowledge from behavioral log data, online content itself contains large amounts of factual knowledge that can be made accessible to automated reasoning. Towards this end, I have built and designed a system [11] which can *automatically* extract massive amounts of knowledge from the semi-supervised content of websites in a robust and efficient fashion by incorporating ideas from wrapper induction and distant supervision.

Another line of work that I pursued tackles the problem of transfer learning between multiple languages. In particular, how can we use methods and data available for well-studied languages (*e.g.,* English) to help tasks in less well-studied languages such as Finnish. To this effect, I have developed algorithms that improve search quality [12, 13] for other languages using Pseudo-Relevance Feedback [14].

Future Research

As I have done over the course of my PhD studies, I plan to continue my research agenda of using behavioral data for improving intelligent systems and for extracting world knowledge. Some of the avenues I hope to explore include:

Representing Knowledge from Behavioral Logs

While my research has focused on extracting knowledge from interaction data for improving learning, an equally critical question is that of how such knowledge can be represented so as to enable easy and efficient learning and inference. Unlike factual knowledge, for which there have been decades of research on representation, behavioral knowledge is very different and can be used for a distinct set of tasks. I believe that *embeddings* are a particularly informative and succinct means of representing such knowledge, if they are endowed with a probabilistic semantic that facilitates reasoning and learning. Embedding models can be used to represent complex, heterogeneous objects in the same space which can allow us to answer rich questions about the similarities between these different objects. For instance, using query log data from search engines, we can look to embed documents, query phrases and user needs all in the same Euclidean space. This in turn would allow us to answer different search and recommendation tasks (beyond just query-document search which was the source of the behavioral data) as well as tackle interesting personalization questions. I foresee a series of rich theoretical and practical questions that need to be tackled in this space including the issues of noise and biases. I also envision a broad set of more domain specific questions, such as the embedding of educational topics and concepts using data from adaptive tutoring systems.

Learning to Proactively Assist on Complex Tasks

Consider a user planning a business trip. To start with, she would need to determine logistics such as travel and accommodation. Once there, she will require further information such as dining locations that she would like as well as convenient commuting options. All put together this is a complex task that would require significant effort from the user, searching and navigating through different resources. Unfortunately, today's information systems lack the tools to help this user. Proactive assistance is a particularly promising option to help users. This requires the learning system to predict what content the user would like to access in the immediate future – for instance, by determining that the user would need dining and commuting suggestions once the user has reached their destination. By correctly predicting this future content we can drastically reduce user effort and improve overall user experience. However given the paucity of context this proactive assistance problem is far more challenging than the *reactive* assistance problems which current information systems address. By leveraging large-scale interaction logs, I hope to come up with robust proactive systems for addressing these challenges.

Innovative applications: Smart Homes, MOOCs, and Beyond

Over the last year I have studied different problems in the education domain as we approach an age where education will further move online, will become adaptive to the student, and will likely increase its reach through lowered costs. All of these trends generate exciting applications and research questions for machine learning with humans in the loop. For example, in addition to problems pertaining to student evaluation at scale, I am also interested in studying means to incentivize students to produce high-quality content (*e.g.,* practice test questions). Furthermore, I conjecture that we can use student interaction data to create personal educational assistants, which can adapt a student's curriculum based on how quickly the student grasps specific concepts, adjust depth based on student interest, and customize tests to minimize testing overhead.

Smart homes are another information system of growing importance. By having connected homes we are able to control different personal devices, automate chores and assist in day-to-day functions. I hope to leverage the rich behavioral data that smart homes provide to help create technologies for these homes of the future. In particular, I am passionate about problems pertaining to energy technologies and hope to use behavioral pattern data to help reduce overall energy consumption.

References

[1] Karthik Raman, Pannaga Shivaswamy, and Thorsten Joachims. Online learning to diversify from implicit feedback. In *KDD*, pages 705–713, 2012.

[2] Karthik Raman, Thorsten Joachims, Pannaga Shivaswamy, and Tobias Schnabel. Stable Coactive Learning via Perturbation. In *ICML*, pages 837–845, 2013.

[3] Karthik Raman, Krysta M. Svore, Ran Gilad-Bachrach, and Chris J. C. Burges. Learning from mistakes: towards a correctable learning algorithm. In *CIKM*, pages 1930–1934, 2012.

[4] Karthik Raman, Paul N. Bennett, and Kevyn Collins-Thompson. Toward whole-session relevance: exploring intrinsic diversity in web search. In *SIGIR*, pages 463–472, 2013.

[5] Karthik Raman, Paul N. Bennett, and Kevyn Collins-Thompson. Understanding intrinsic diversity in web search: Improving whole-session relevance. *ACM Trans. Inf. Syst.*, 32(4):20:1–20:45, October 2014.

[6] Karthik Raman and Thorsten Joachims. Learning Socially Optimal Information Systems from Egoistic Users. In *ECML/PKDD (2)*, pages 128–144, 2013.

[7] Karthik Raman, Thorsten Joachims, and Pannaga Shivaswamy. Structured learning of two-level dynamic rankings. In *CIKM*, pages 291–296, 2011.

[8] Karthik Raman and Thorsten Joachims. Methods for ordinal peer grading. In *Proceedings of the 20th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, KDD '14, pages 1037–1046, New York, NY, USA, 2014. ACM.

[9] Karthik Raman and Thorsten Joachims. Bayesian methods for ordinal peer grading. In *ACM Learning at Scale Conference*, 2015.

[10] Karthik Raman, Adith Swaminathan, Thorsten Joachims, and Johannes Gehrke. Beyond myopic inference in big data pipelines. In *KDD*, pages 86–94, 2013.

[11] Karthik Raman, Joe Dahlquist, Jeffrey Dalton, Evgeniy Gabriloivch, Kevin Murphy, and Wei Zhang. RAKE: Robust Automatic Knowledge Extraction from semi-structured web pages. In Submission to WWW 2015.

[12] Manoj K. Chinnakotla, Karthik Raman, and Pushpak Bhattacharyya. Multilingual PRF: English lends a helping hand. In *SIGIR*, pages 659–666, 2010.

[13] Manoj K. Chinnakotla, Karthik Raman, and Pushpak Bhattacharyya. Multilingual pseudo-relevance feedback: performance study of assisting languages. In *ACL*, pages 1346–1356, 2010.

[14] Karthik Raman, Raghavendra Udupa, Pushpak Bhattacharya, and Abhijit Bhole. On improving pseudo-relevance feedback using pseudo-irrelevant documents. In *ECIR*, pages 573–576, 2010.