with universal choice sets in Figure 1). There is no methodology for learning determinantal point processes with variable choice sets, so we do not compare against them here.

Figure 3 shows the mean per-choice likelihood improvements over the separable model. In both datasets, our model achieves over 5% likelihood gain with $|H| < 10$. With YcItems, we see the same pattern in likelihood gains as for the datasets with universal choice sets: a rapid increase for the first several corrections and then a leveling of the gains. For this dataset, $|H|$ is less than 1% the number of items, and we achieve substantial likelihood gains with an extremely sparse model.

5 RELATED WORK

There are variable-size choice models for universal choice sets based on pairwise interactions and conditional distributions [26], correlated random utility errors [17], or a priori knowledge of utilities [11, 35]. Our model makes no assumptions on pairwise interactions and can handle arbitrary-order interactions. Also, the datasets used in our experiments contain orders of magnitude more items than experiments for pairwise interaction models in market basket data [5, 9]. Set prediction functions in neural networks [22, 37] and mutli-label classification methods more broadly [23, 34] are also relevant to the subset prediction problem. However, these methods are designed to predict a set of labels from features (e.g., predict several tags of an image), whereas our experiments predict new subsets given previous subset selections. The recently developed set embedding model offers a statistical approach to the subset choice problem with universal choice sets [25]. Unlike our models, neural networks and the embedding models are not known to carry a random utility maximization interpretation.

Another variant of subset choice is approval voting, where an individual selects all candidates that she approves [10, 24, 29]. In this case, there is an implicit assumption that only one alternative will ultimately be selected (only one candidate will win the election), whereas we deal with subsets that give complementary utility.

Lastly, our work fits into the context of several recent analyses on effectively learning discrete choice models [4, 16, 18, 19, 21, 38] as well as applications of discrete choice models to user behavior on the Web [7, 30, 36].

6 DISCUSSION

We developed a general random utility model for how individuals choose a subset of items from a given choice set and analyzed its structure in two contexts: (i) the choice set is universal and is the same for all selections and (ii) the choice set varies. In both cases, we prove that after identifying subsets receiving corrective probability within our model, we can efficiently find the optimal model parameters. However, we also showed that finding the best set of subsets to receive corrections is NP-hard. Approximation algorithms for coping with this issue are a direction for future work. Nevertheless, our sparse model provides substantial likelihood improvements in subset prediction compared to competing baselines.

REFERENCES

[1] Rakesh Agrawal and Ramakrishnan Srikant. 1994. Fast algorithms for mining association rules. In *VLDB*.

[2] Moshe E Ben-Akiva and Steven R Lerman. 1985. *Discrete Choice Analysis: Theory and Application to Travel Demand.* MIT Press.

[3] Austin R. Benson, Ravi Kumar, and Andrew Tomkins. 2016. Modeling User Consumption Sequences. In *WWW*. 519–529.

[4] Austin R. Benson, Ravi Kumar, and Andrew Tomkins. 2016. On the Relevance of Irrelevant Alternatives. In *WWW*. 963–973.

[5] Yasemin Boztuğ and Lutz Hildebrandt. 2006. A market basket analysis conducted with a multivariate logit model. In *From Data and Information Analysis to Knowledge Engineering*. Springer, 558–565.

[6] O. Celma. 2010. *Music Recommendation and Discovery in the Long Tail.* Springer.

[7] Shuo Chen and Thorsten Joachims. 2016. Modeling intransitivity in matchup and comparison data. In *WSDM*. 227–236.

[8] David R Cox. 1958. The regression analysis of binary sequences. *Journal of the Royal Statistical Society. Series B (Methodological)* (1958), 215–242.

[9] Katrin Dippold and Harald Hruschka. 2013. Variable selection for market basket analysis. *Computational Statistics* 28, 2 (2013), 519–539.

[10] J-Cl Falmagne and Michael Regenwetter. 1996. A random utility model for approval voting. *Journal of Mathematical Psychology* 40, 2 (1996), 152–159.

[11] Peter C Fishburn and Irving H LaValle. 1996. Binary interactions and subset choice. *European Journal of Operational Research* 92, 1 (1996), 182–192.

[12] Jennifer A Gillenwater, Alex Kulesza, Emily Fox, and Ben Taskar. 2014. Expectation-maximization for learning determinantal point processes. In *NIPS*.

[13] Nick Hanley, Robert E Wright, and Vic Adamowicz. 1998. Using choice experiments to value the environment. *Environ. Resour. Econ* 11, 3-4 (1998), 413–428.

[14] Instacart. 2017. The Instacart Online Grocery Shopping Dataset. https://www.instacart.com/datasets/grocery-shopping-2017. (2017).

[15] Alex Kulesza and Ben Taskar. 2012. Determinantal point processes for machine learning. *Foundations and Trends in Machine Learning* 5, 2–3 (2012), 123–286.

[16] Ravi Kumar, Andrew Tomkins, Sergei Vassilvitskii, and Erik Vee. 2015. Inverting a steady-state. In *WSDM*. 359–368.

[17] Puneet Manchanda, Asim Ansari, and Sunil Gupta. 1999. The "shopping basket": a model for multicategory purchase incidence decisions. *Marketing Science* 18, 2 (1999), 95–114.

[18] Lucas Maystre and Matthias Grossglauser. 2015. Fast and accurate inference of Plackett–Luce models. In *NIPS*.

[19] Lucas Maystre and Matthias Grossglauser. 2017. ChoiceRank: Identifying Preferences from Node Traffic in Networks. In *ICML*.

[20] Jong Soo Park, Ming-Syan Chen, and Philip S. Yu. 1995. An Effective Hash-based Algorithm for Mining Association Rules. In *SIGMOD*. 175–186.

[21] Stephen Ragain and Johan Ugander. 2016. Pairwise Choice Markov Chains. In *NIPS*.

[22] Siamak Ravanbakhsh, Jeff Schneider, and Barnabas Poczos. 2017. Deep learning with sets and point clouds. In *ICLR Workshop Track*.

[23] Jesse Read, Bernhard Pfahringer, Geoff Holmes, and Eibe Frank. 2011. Classifier chains for multi-label classification. *Machine Learning* 85, 3 (2011), 333–359.

[24] Michel Regenwetter and Bernard Grofman. 1998. Choosing subsets: a size-independent probabilistic model and the quest for a social welfare ordering. *Social Choice and Welfare* 15, 3 (1998), 423–443.

[25] Maja Rudolph, Francisco Ruiz, Stephan Mandt, and David Blei. 2016. Exponential family embeddings. In *NIPS*.

[26] Gary J Russell and Ann Petersen. 2000. Analysis of cross category dependence in market basket selection. *Journal of Retailing* 76, 3 (2000), 367–392.

[27] Ashoka Savasere, Edward Omiecinski, and Shamkant B Navathe. 1995. An Efficient Algorithm for Mining Association Rules in Large Databases. In *VLDB*.

[28] J Ben Schafer, Joseph Konstan, and John Riedl. 1999. Recommender systems in e-commerce. In *EC*. 158–166.

[29] Nihar Shah, Dengyong Zhou, and Yuval Peres. 2015. Approval voting and incentives in crowdsourcing. In *ICML*.

[30] Or Sheffet, Nina Mishra, and Samuel Ieong. 2012. Predicting Consumer Behavior in Commerce Search. In *ICML*.

[31] Stefan Stremersch and Gerard J Tellis. 2002. Strategic bundling of products and prices: A new synthesis for marketing. *Journal of Marketing* 66, 1 (2002), 55–72.

[32] Kenneth Train. 1986. *Qualitative Choice Analysis: Theory, Econometrics, and an Application to Automobile Demand.* Vol. 10. MIT press.

[33] Kenneth E Train. 2009. *Discrete Choice Methods with Simulation.* Cambridge University Press.

[34] Grigorios Tsoumakas and Ioannis Vlahavas. 2007. Random k-labelsets: An ensemble method for multilabel classification. In *ECML*. 406–417.

[35] Iris van Rooij, Ulrike Stege, and Helena Kadlec. 2005. Sources of complexity in subset choice. *Journal of Mathematical Psychology* 49, 2 (2005), 160–187.

[36] Shuang-Hong Yang, Bo Long, Alexander J Smola, Hongyuan Zha, and Zhaohui Zheng. 2011. Collaborative competitive filtering: learning recommender using context of user choice. In *SIGIR*. 295–304.

[37] Manzil Zaheer, Satwik Kottur, Siamak Ravanbakhsh, Barnabas Poczos, Ruslan Salakhutdinov, and Alexander Smola. 2017. Deep Sets. In *NIPS*.

[38] Danqing Zhang, Kimon Fountoulakis, Junyu Cao, Mogeng Yin, Michael Mahoney, and Alexei Pozdnoukhov. 2017. Social Discrete Choice Models. *arXiv* (2017).