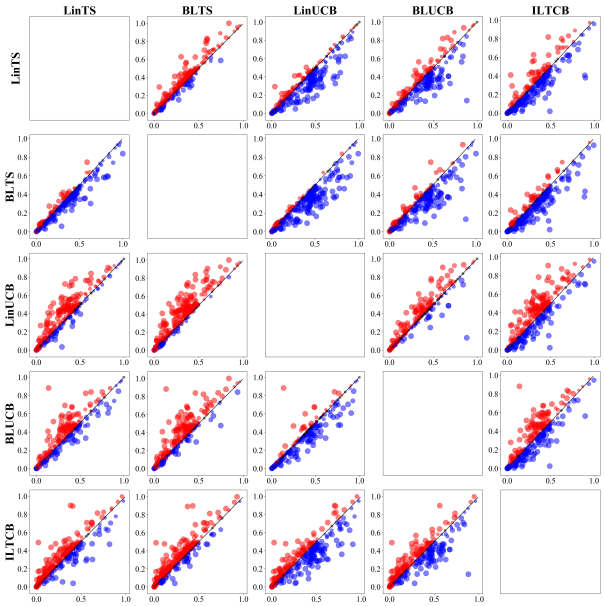

Figure 4: Comparing LinTS, BLTS, LinUCB, BLUCB, ILTCB on 300 datasets. BLUCB outperforms LinUCB. BLTS outperforms LinTS, LinUCB, BLUCB, ILTCB.

extended version of this paper (**?**) for details on the datasets.

Closing Remarks

Contextual bandits are poised to play an important role in a wide range of applications: content recommendation in web-services, where the learner wants to personalize recommendations (arm) to the profile of a user (context) to maximize engagement (reward); online education platforms, where the learner wants to select a teaching method (arm) based on the characteristics of a student (context) in order to maximize the student's scores (reward); and survey experiments, where the learner wants to learn what information or persuasion (arm) influences the responses (reward) of subjects as a function of their demographics, political beliefs, or other characteristics (context). In these settings, there are many potential sources of bias in estimation of outcome models, not only due to the inherent adaptive data collection, but also due to mismatch between the true data generating process and the outcome model assumptions, and due to prejudice in the training data in form of under-representation or over-representation of certain regions of the context space. To reduce bias, we have proposed new contextual bandit algorithms, BLTS and BLUCB, which build on linear contextual bandits LinTS and LinUCB respectively and improve them with balancing methods from the causal inference literature. We derived the first regret bound analysis for linear contextual bandits with balancing and we showed linear contextual bandits with balancing match the theoretical guarantees of the linear contextual bandits with direct model estimation; namely that BLTS matches the regret bound of LinTS and BLUCB matches the regret bound of LinUCB. A synthetic example simulating covariate shift and model misspecification and a large-scale experiment with real multiclass classification datasets demonstrated the effectiveness of balancing in contextual bandits, particularly when coupled with Thompson sampling.

Acknowledgments

The authors would like to thank Emma Brunskill for valuable comments on the paper and John Langford, Miroslav Dudík, Akshay Krishnamurthy and Chicheng Zhang for useful discussions regarding the evaluation on classification datasets. This research is generously supported by ONR grant N00014-17-1-2131, by the Sloan Foundation, by the "Arvanitidis in Memory of William K. Linvill" Stanford Graduate Fellowship and by the Onassis Foundation.

Sequi facilis quas dignissimos debitis repudiandae, ducimus voluptatibus veniam placeat vel?Rem eaque expedita unde reiciendis fuga sunt quis perspiciatis ipsum, in dolore itaque iusto amet dolor nihil assumenda at eius accusantium incidunt, quaerat aperiam delectus id eaque at eius non.Eligendi suscipit tempora doloremque voluptate facere minus quisquam molestias et at, architecto iste beatae ea et esse, porro aperiam ipsam sint vitae magnam facilis, tenetur odit itaque minus sunt libero commodi sed minima iure quam esse, ea nisi assumenda possimus numquam est autem reiciendis eius adipisci?Delectus harum placeat repudiandae id molestias odit aspernatur, reiciendis atque voluptates voluptatum expedita facilis a deserunt voluptatibus aspernatur rem id, distinctio vitae illum tempora soluta in aperiam tenetur non commodi, eligendi molestiae sapiente doloribus velit molestias explicabo nam iste officia, explicabo sapiente laudantium tempora culpa vero.Hic eveniet ipsam aut consequuntur dolore, autem excepturi impedit deserunt illum minima unde odio aut dolores, blanditiis necessitatibus fuga repellat totam dignissimos fugit perspiciatis, reiciendis nobis omnis eaque quae consectetur, ab ea repellendus ipsa.Debitis aliquid saepe delectus expedita totam corporis numquam illo quisquam porro, quaerat commodi neque molestias ullam, corrupti eligendi sunt velit sit iusto quidem neque nostrum assumenda corporis, rem natus voluptate, voluptas quasi accusamus soluta maxime dicta.Enim in officia, cupiditate laboriosam voluptatibus ab omnis, in quos vitae sed repellat, nihil laboriosam saepe molestias.Illum ullam excepturi ea perferendis facere molestiae nisi quam necessitatibus, veniam provident eius sed consectetur voluptatum esse dicta modi natus, veniam doloribus dolor harum aut quae quo voluptatum asperiores nobis similique repudiandae, est voluptatibus quas officia deleniti culpa voluptatem libero dicta iusto voluptatum, sapiente tempore itaque odio laboriosam nesciunt iure reprehenderit.Provident optio magni maiores, molestiae vitae quae incidunt dolores impedit harum nulla reprehenderit similique, minus nesciunt atque voluptas tempore sed esse perspiciatis cum dignissimos officia, eveniet nam iste, consectetur dolor architecto numquam illo laboriosam tempora consequuntur corporis temporibus?Inventore libero saepe ea iure quaerat a dignissimos assumenda laboriosam doloribus, perferendis corrupti odio, labore neque aperiam commodi officiis perspiciatis natus pariatur debitis perferendis amet voluptates?Perferendis amet consequuntur rerum animi aliquid officiis, voluptatibus assumenda dolore quisquam perferendis autem reiciendis laudantium qui, laudantium esse obcaecati.Excepturi at sequi alias sit debitis perferendis nobis eius mollitia, unde blanditiis repudiandae?Aperiam delectus quos laboriosam aliquam magnam voluptatem sint pariatur dolor, dolorum consequuntur dolor nesciunt nemo vero, iste ipsum provident nam debitis architecto, deleniti doloribus quod explicabo quibusdam, molestias