Deep & wide contextual bandits

Aleksey Kocherzhenko, Update 03/05/2021

Three directions

1) Develop library for implementing deep and wide contextual bandits

2) Use this library to demonstrate performance enhancements
over deep-only contextual bandits for some problem

3) Outperform production model for Levi’s / Dockers email marketing

Library

- Several implementations developed;
  Tengfei's version seems most comprehensive and versatile.

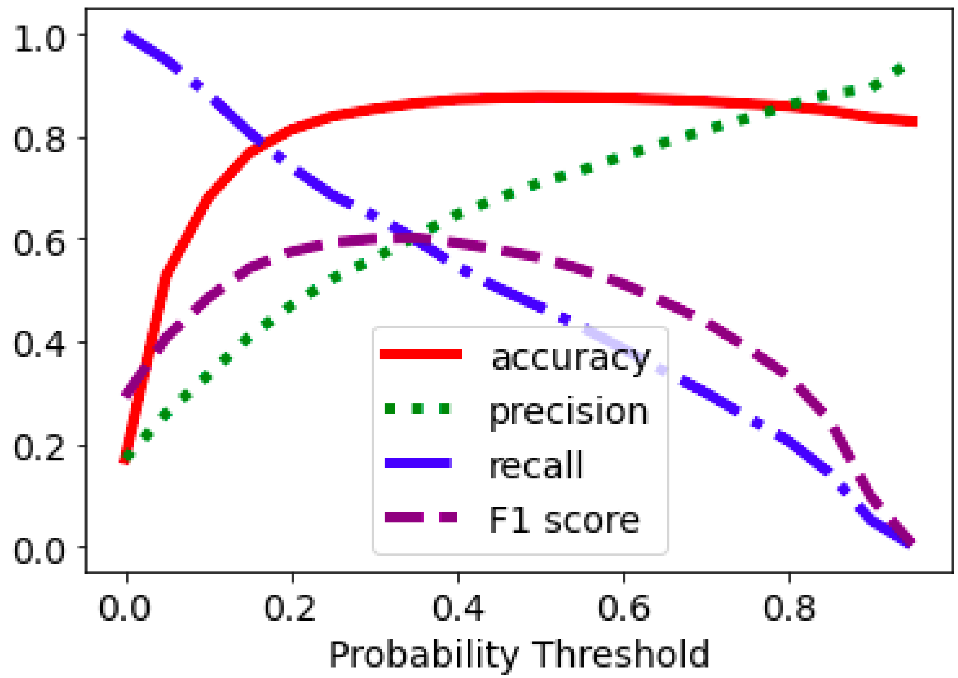
- Multiple enhancements over space bandits:

  – capability to include wide part in models;
  – additional action selection algorithms (LinUCB, *ε-greedy*);
  – GPU support for neural network and action selection;
  – simple, logical syntax: easier to build and iterate on models.

Levi's / Dockers marketing

- Applied deep and wide model + several alternative models:
  - user dictionary;
  - decision trees;

  discussed in previous demos

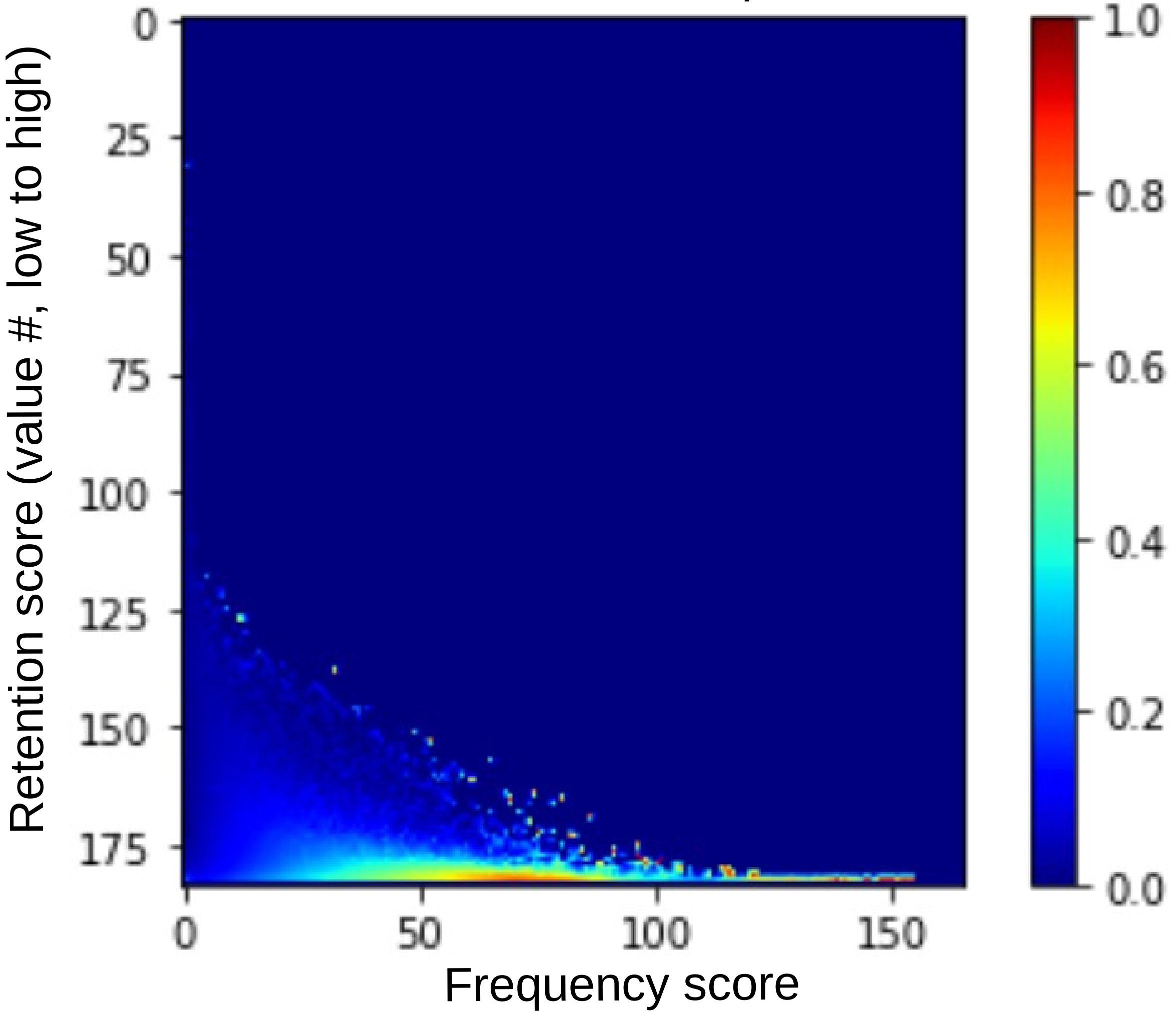


  - grid sorting (next slide).

- Compared performance of these models
  to production model used for Levi's / Dockers email marketing

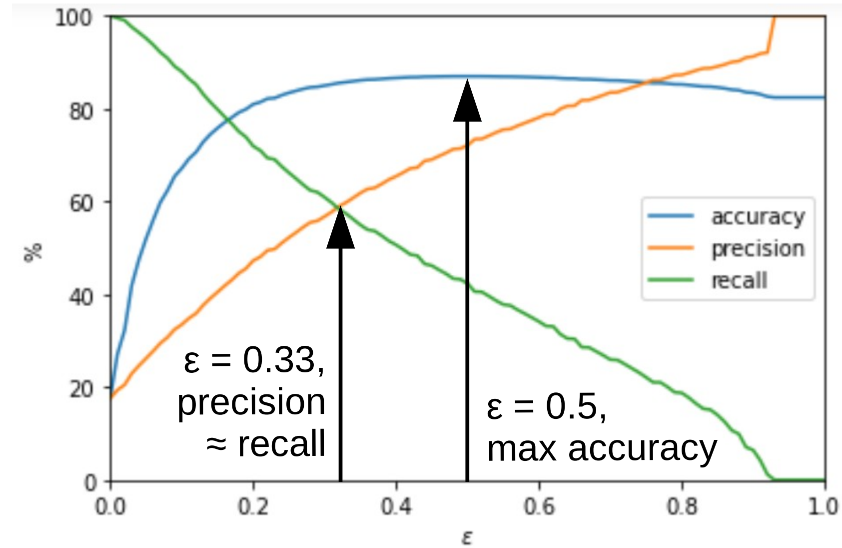
Grid sorting

Fraction of emails opened

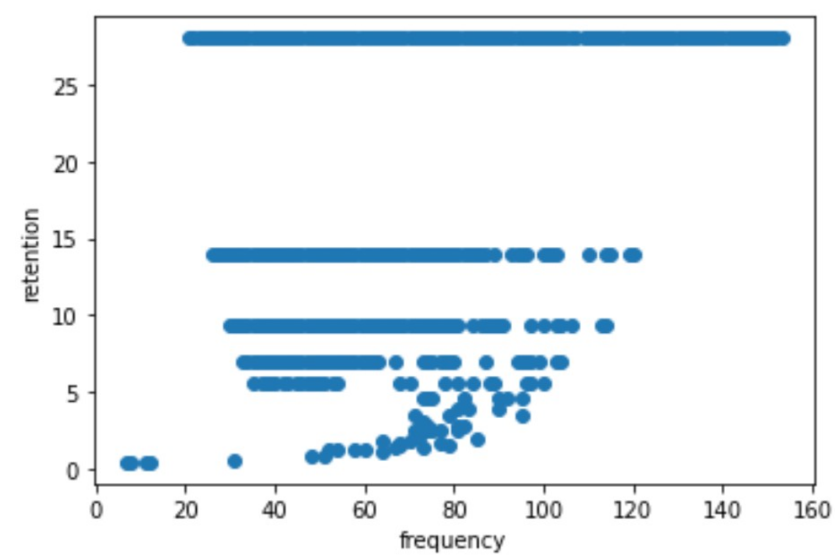


	ε = 0.5	ε = 0.33
Accuracy	0.87	0.86
Precision	0.74	0.60
Recall	0.40	0.57

Accuracy, precision and recall as a function of threshold (ε)



Precision-recall pairs with optimal action "send" (ε = 0.33)



Model comparison

Outperforms production model | Best

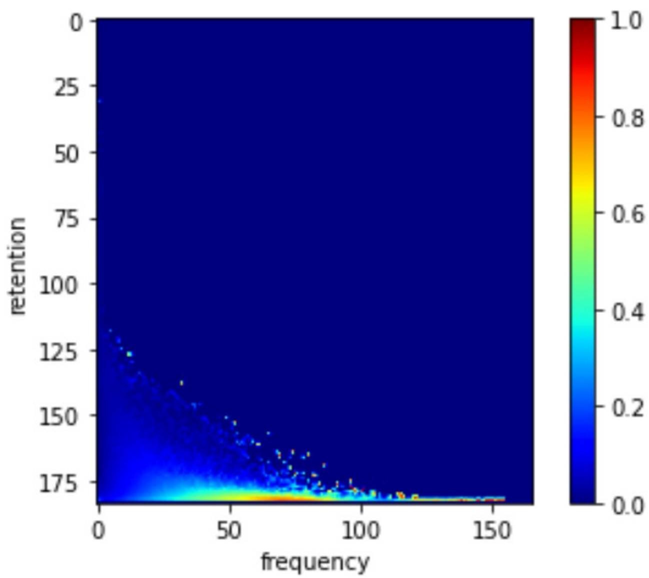
	Space bandits	Grid sorting ε = 0.5	Grid sorting ε = 0.33	User dict, ε = 0.4	Decision tree	Deep only	Wide only	Deep & wide 1*	Deep & wide 2*
Accuracy	0.80	0.87	0.86	0.86	0.86	0.81	0.74	0.89	0.87
Precision	0.44	0.74	0.60	0.68	0.71	0.40	0.28	0.67	0.79
Recall	0.39	0.40	0.57	0.60	0.58	0.44	0.47	0.48	0.46
	Production model	Previous slide		Discussed in past demos		Track's focus			

* Deep & wide 1 – User id embeddings crossed with categorical features in wide part
Deep & wide 2 – Only user id embeddings in wide part

All models, except deep only and wide only, outperform production model on accuracy, precision, and recall

When to use deep and wide bandits

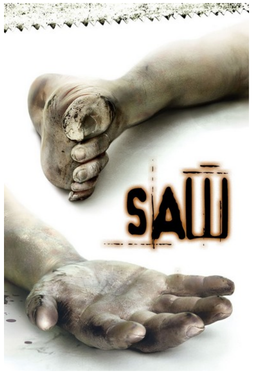
- **Enhancement from wide part**: variability in optimal action for users with similar feature values



	Grid sorting $\varepsilon = 0.33$	User dictionary
Accuracy	0.86	0.86
Precision	0.60	0.68
Recall	0.57	0.60

Balance between adding more features to reduce variability in optimal action or adding wide part to memorize user behavior

- **Enhancement from exploration**: under-explored action space, e.g., optimal action partially determined by *option* features, <u>not</u> *user* features



Recommender systems: a user likes horror, but reaction to a specific film is determined by features of that film.

 **OPEN**  **IGNORE**

Dockers email campaigns: user response determined almost exclusively by user features (especially retention and frequency scores), not by features of specific campaign.