Feature interactions

Example: banner selection

•••	category_ad	category_site	•••	is_clicked
•••	auto_part	game_news	•••	0
•••	music_tickets	music_news	••	1
•••	mobile_phones	auto_blog	•••	0

Example: banner selection

•••	category_ad	category_site	•••	is_clicked
•••	auto_part	game_news	•••	0
•••	music_tickets	music_news	••	1
•••	mobile_phones	auto_blog	•••	0

•••	ad_site	•••	is_clicked
•••	auto_part game_news	•••	0
•••	music_tikets music_news	••	1
•••	mobile_phones auto_blog	•••	0

f1	f2
А	Х
В	Y
В	Ζ
А	Ζ

f1	f2		f_join		A X	B Y	BZ	A Z
А	Х	Join	A X	OneHot (f_join)	1			
В	Y		B Y			1		
В	Ζ	F	B Z	٢			1	
А	Ζ		A Z					1





Frequent operations for feature interaction

- Multiplication
- Sum
- Diff
- Division

Practical Notes

- We have a lot of possible interactions N*N for N features.
 - a. Even more if use several types in interactions
- Need to reduce its' number
 - a. Dimensionality reduction
 - b. Feature selection

Example of interaction generation pipeline



Interactions' order

- We looked at 2nd order interactions.
- Such approach can be generalized for higher orders.
- It is hard to do generation and selection automatically.
- Manual building of high-order interactions is some kind of art.

Extract features from DT



Extract features from DT



Extract features from DT



How to use it

In sklearn:

```
tree_model.apply()
```

In xgboost:

booster.predict(pred_leaf=True)

Conclusion

- We looked at ways to build an interaction of categorical attributes
- Extended this approach to real-valued features
- Learn how to extract features via decision trees