



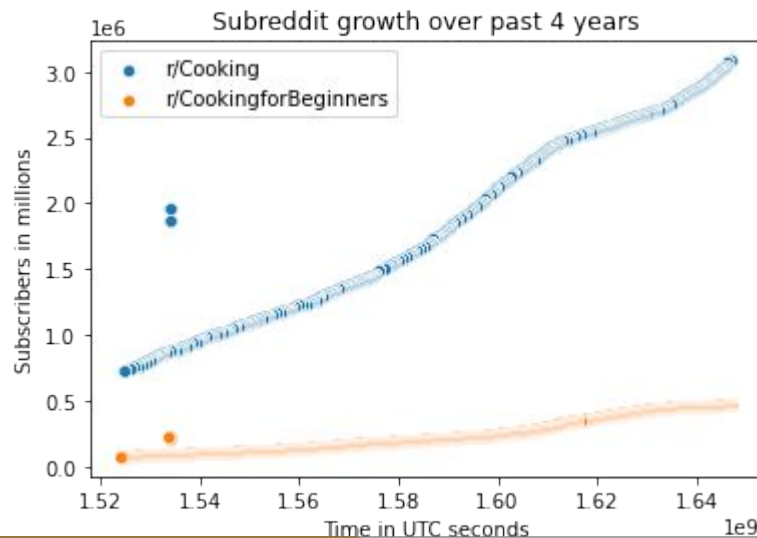
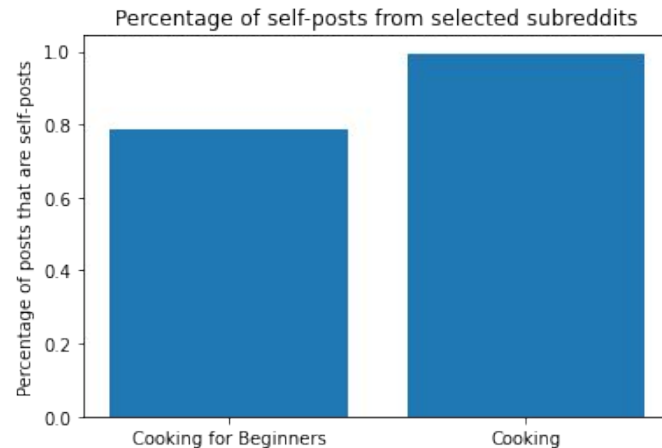
Separating the chefs from the chaff

Comparison and classification of the r/cooking and r/cookingforbeginners subreddits
Karthik Nambiar



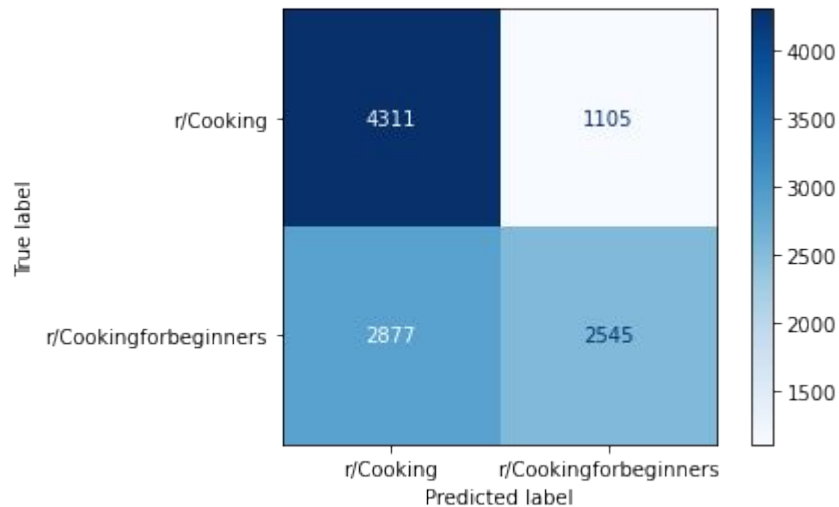
Introduction

- 4 years of posts
- Self-Text and Post Titles
- Steady growth/fairly active
- Mostly self-posts



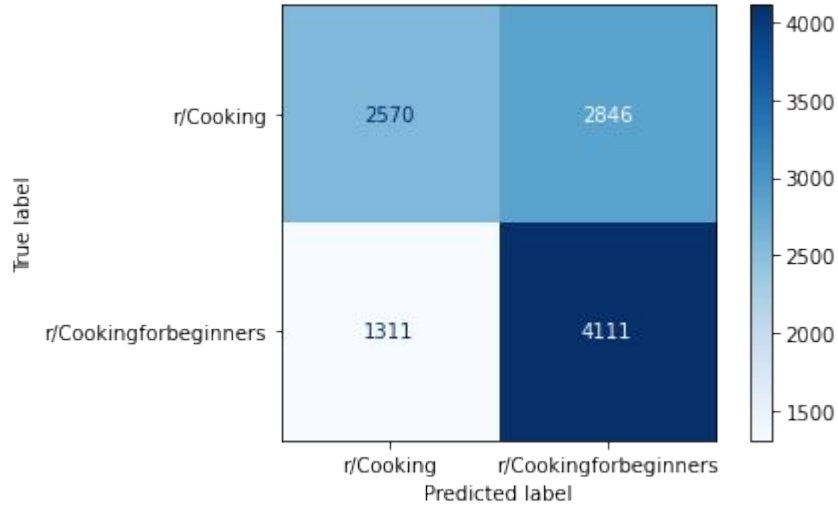
Self-post analysis

- Subreddits are similar
- Different classifiers optimize on different classes
- Combination helps patch holes

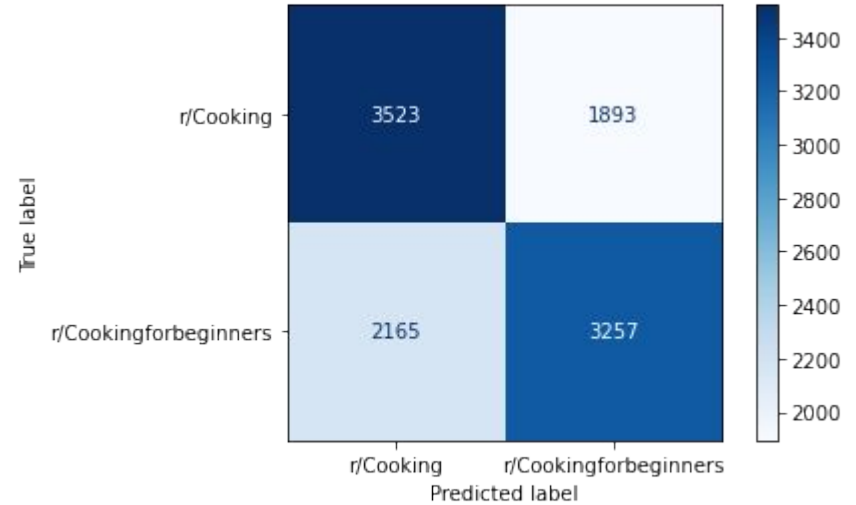


Extra Forests Classifier

Where did the individual models do well?



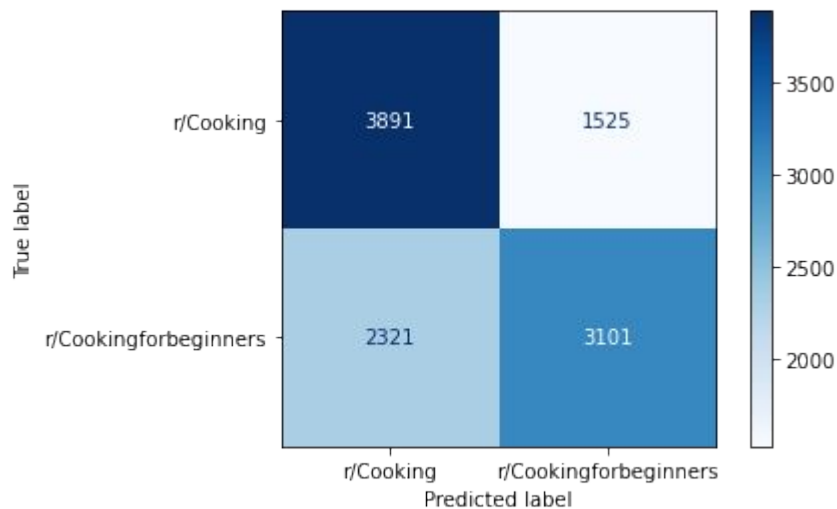
Logistic Regression



Random Forest Classifier

3 heads are better than 1

- 3 other models between 61-62% accuracy
- Joint Model hits 64.5% accuracy
- More even predictions



Where did it fail/succeed?

	ft_imp		ft_imp		ft_imp
removed	0.029470	removed	0.014796	cookingforbeginners	3.615796
and	0.014833	to	0.004962	beginners	2.457308
the	0.013500	of	0.004900	packets	1.696562
to	0.013281	and	0.004802	myweekendplan	1.692092
of	0.011026	it	0.004711	beginner	1.552283
it	0.010756	the	0.004654	unsweetened	1.467371
for	0.009200	for	0.004434	basics	1.387838
in	0.008400	is	0.004352	optional	1.373800
is	0.007713	in	0.004322	thickness	1.328635
with	0.007521	with	0.004244	takeout	1.318424

- Post being removed is important
- Lots of helper words
- Cares about phrases like 'cooking for beginners'

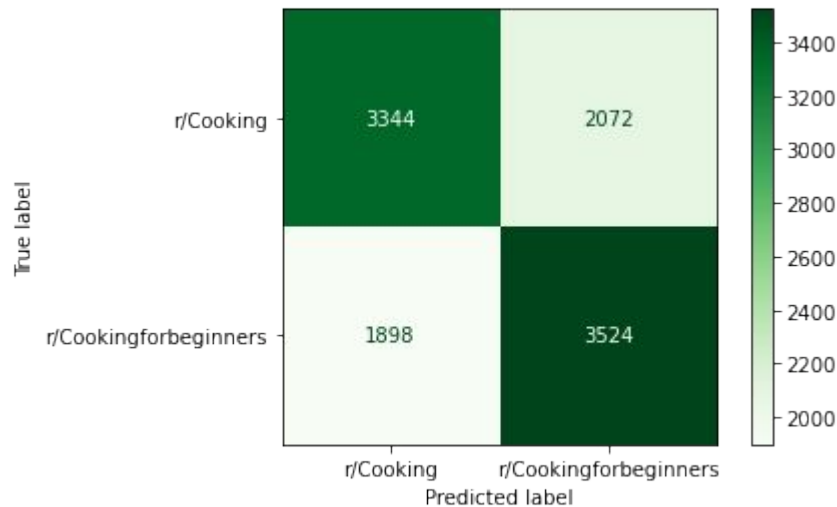
Where did it fail/succeed?

Self-text	All models predicted	But actually...
This is really sad but I don't know how to turn on a stove there is no high or lite	r/CookingforBeginners	r/Cooking
Hi all I have been getting better and better at cooking. I can make some pasta dishes and a few different things here and there. I'm looking for some new ideas. Has anyone got any solid meal lists that everyone should know how to make?	r/Cooking	r/CookingforBeginners

Lesson: There are lots of subtleties in subreddits, and some posts would fit easily in either

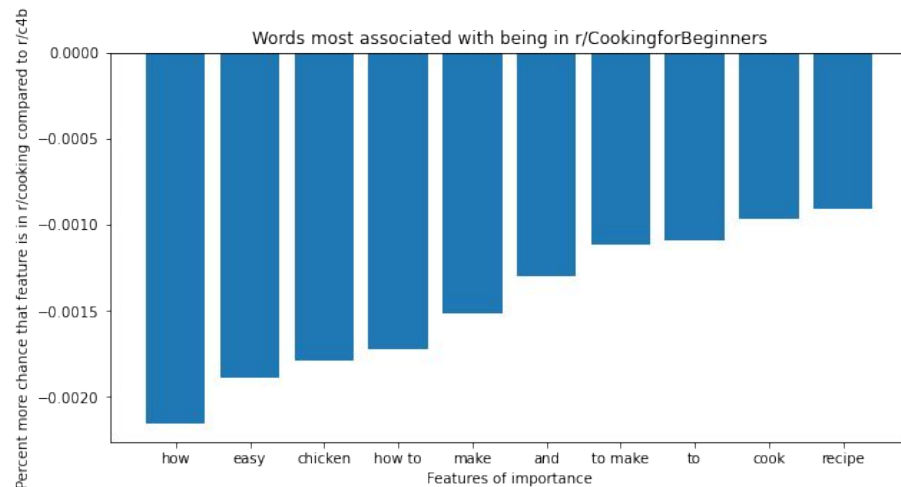
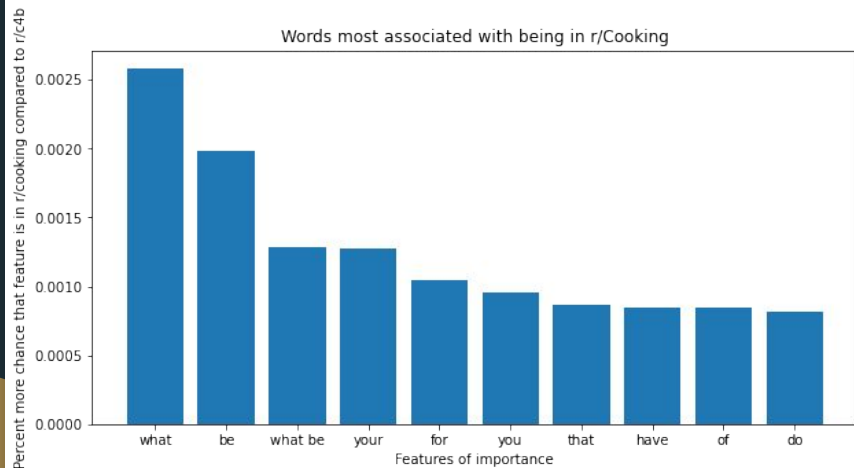
Title

- 3 models- Multinomial NB, Random Forest, Extra Forest
- No odd prediction patterns
- Ensemble outperformed models (~63% accuracy)



Combined Model

Failures/Successes



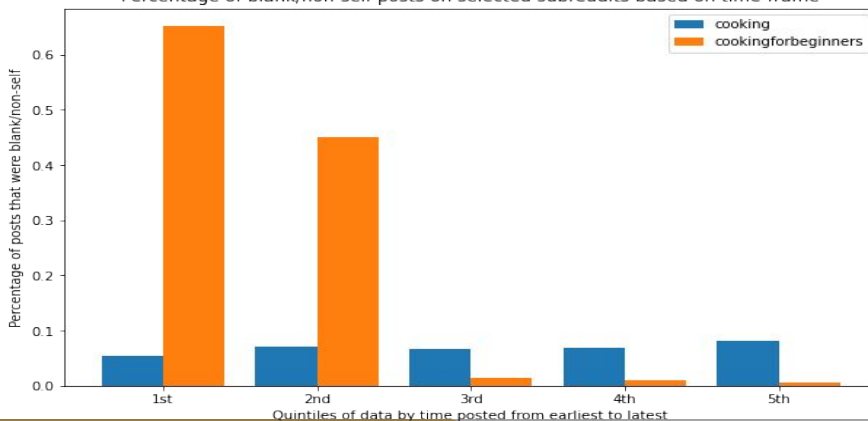
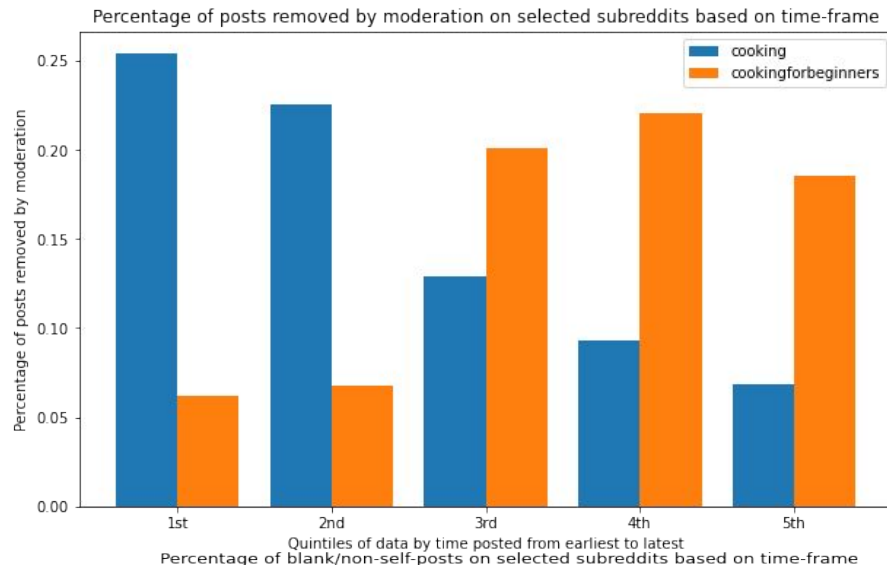
- Other classifiers also found 'chicken', 'easy', 'how to', 'recipe', 'what' useful

Failures/Successes

Title	All models predicted	But actually
How does one actually make beef gravy?	r/CookingforBeginners	r/Cooking
Been trying to perfect my Asian noodle soups....	r/Cooking	r/CookingforBeginners

Future Exploration

- More precise lemmas
- Time series comparison
- Subreddit comparable over the years?
- Comments?



Sources

- <https://github.com/pushshift/api>
- [reddit.com/r/cooking](https://www.reddit.com/r/cooking)
- [reddit.com/r/cookingforbeginners](https://www.reddit.com/r/cookingforbeginners)

Thank you! Questions?