

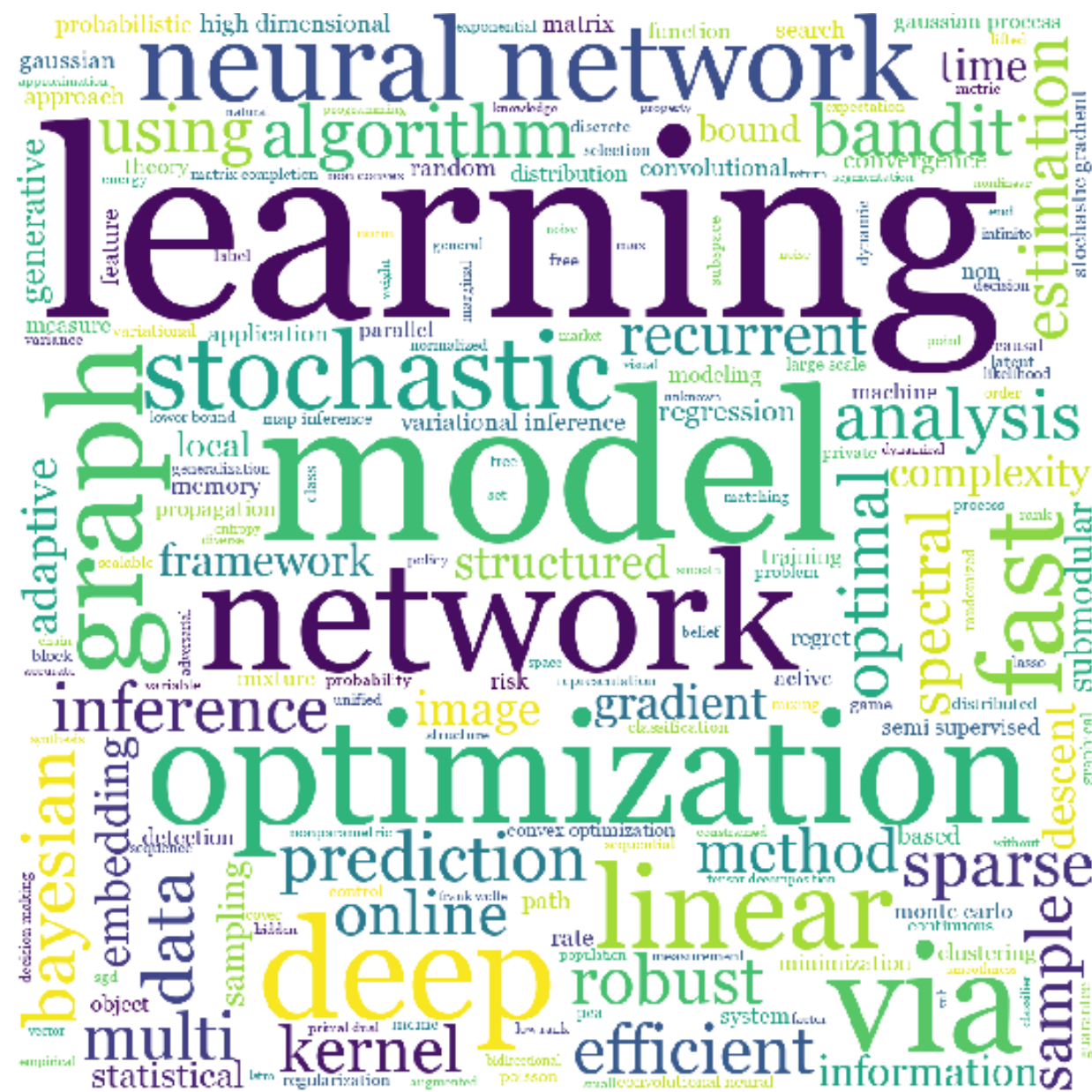
# Визуализация данных

## NIPS 2015 Papers

Каюмов Эмиль, 517  
ММП ВМК МГУ

Курс «Прикладные задачи анализа данных»

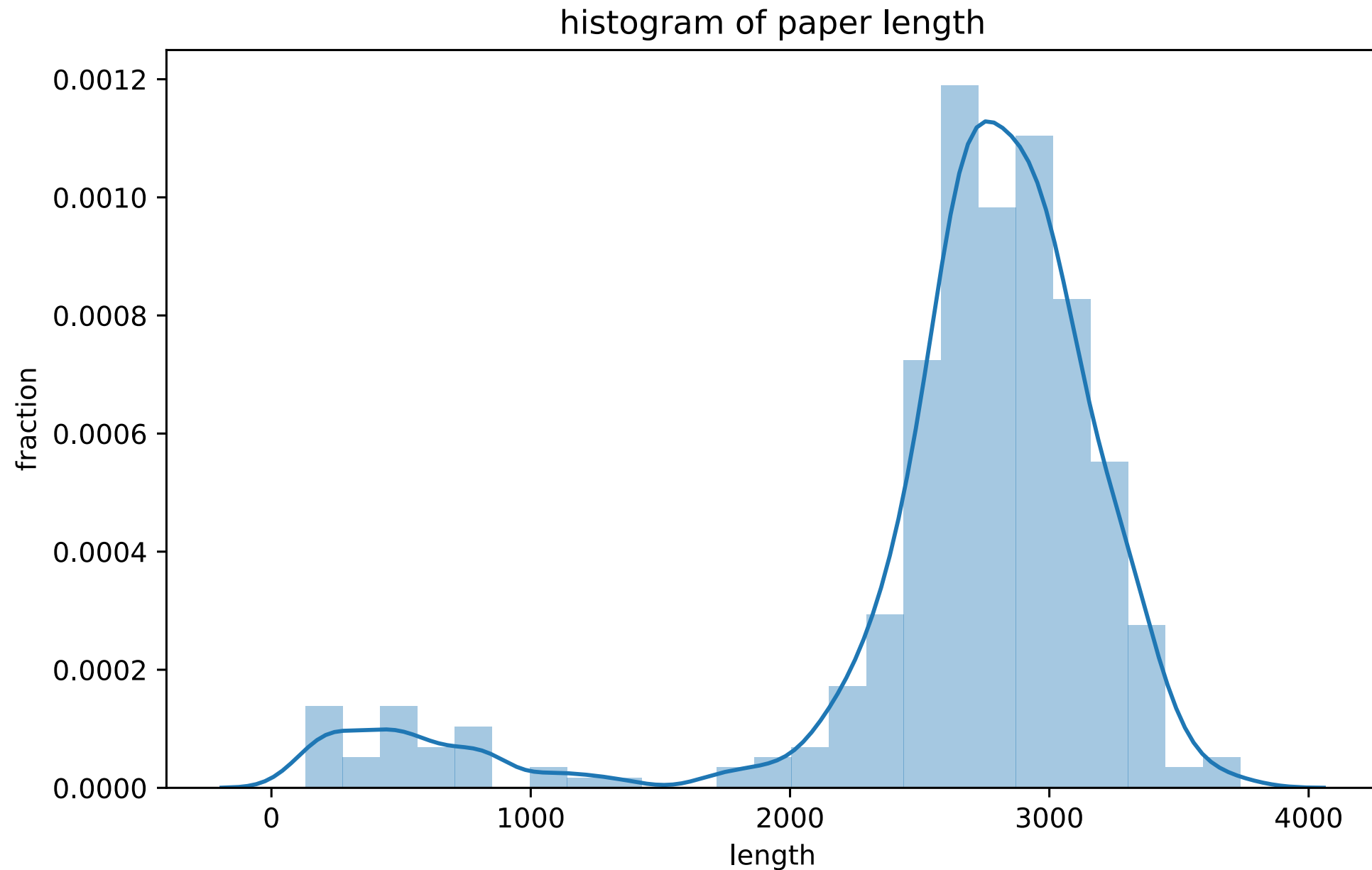
# Популярность слов в заголовках статей



## Популярные слова без стоп-слов и после лемматизации

# Длины статей

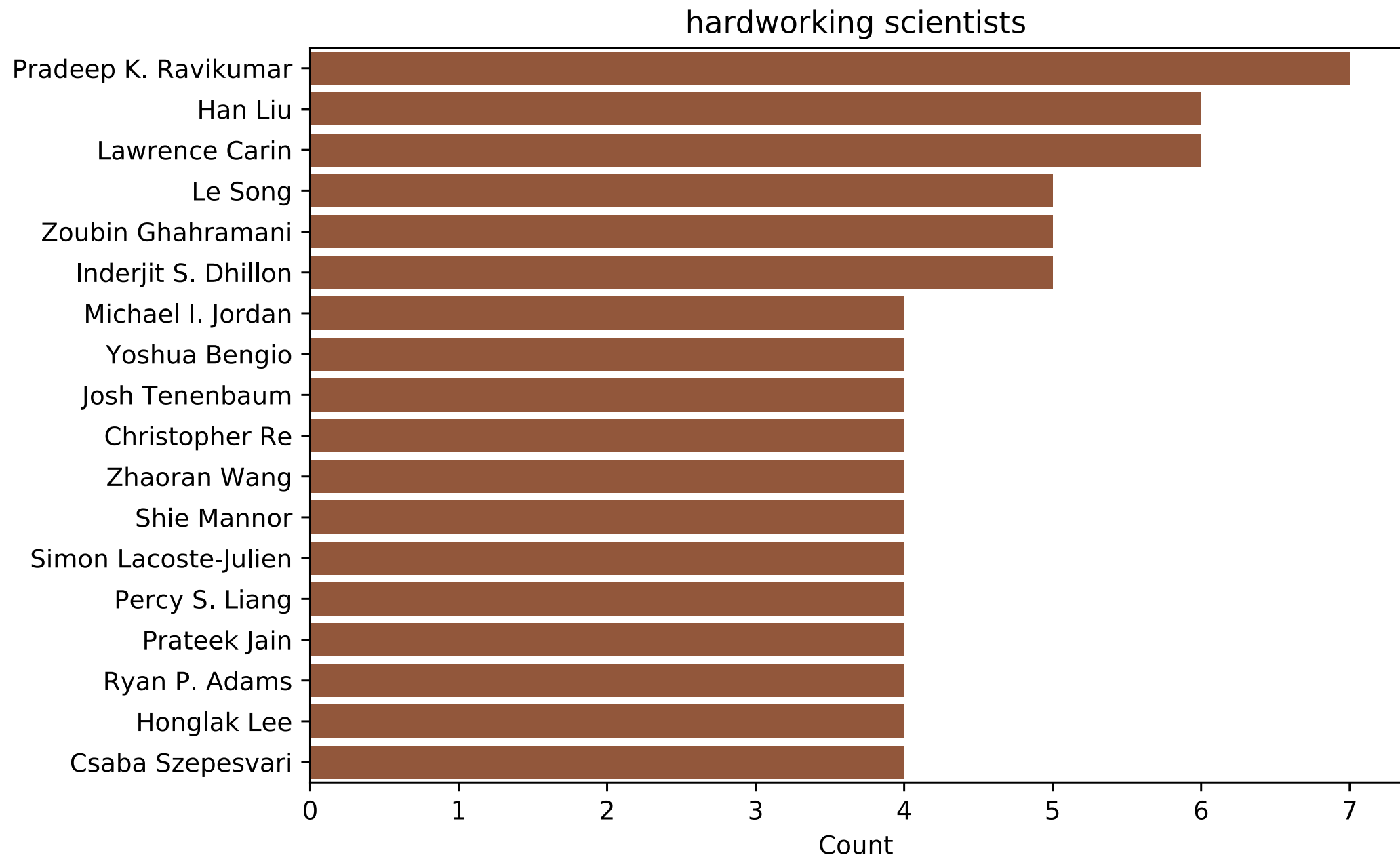
Статьи довольно похожи по объёму



Количество слов в статье без учёта ссылок и стоп-слов

# АКТИВНЫЕ авторы

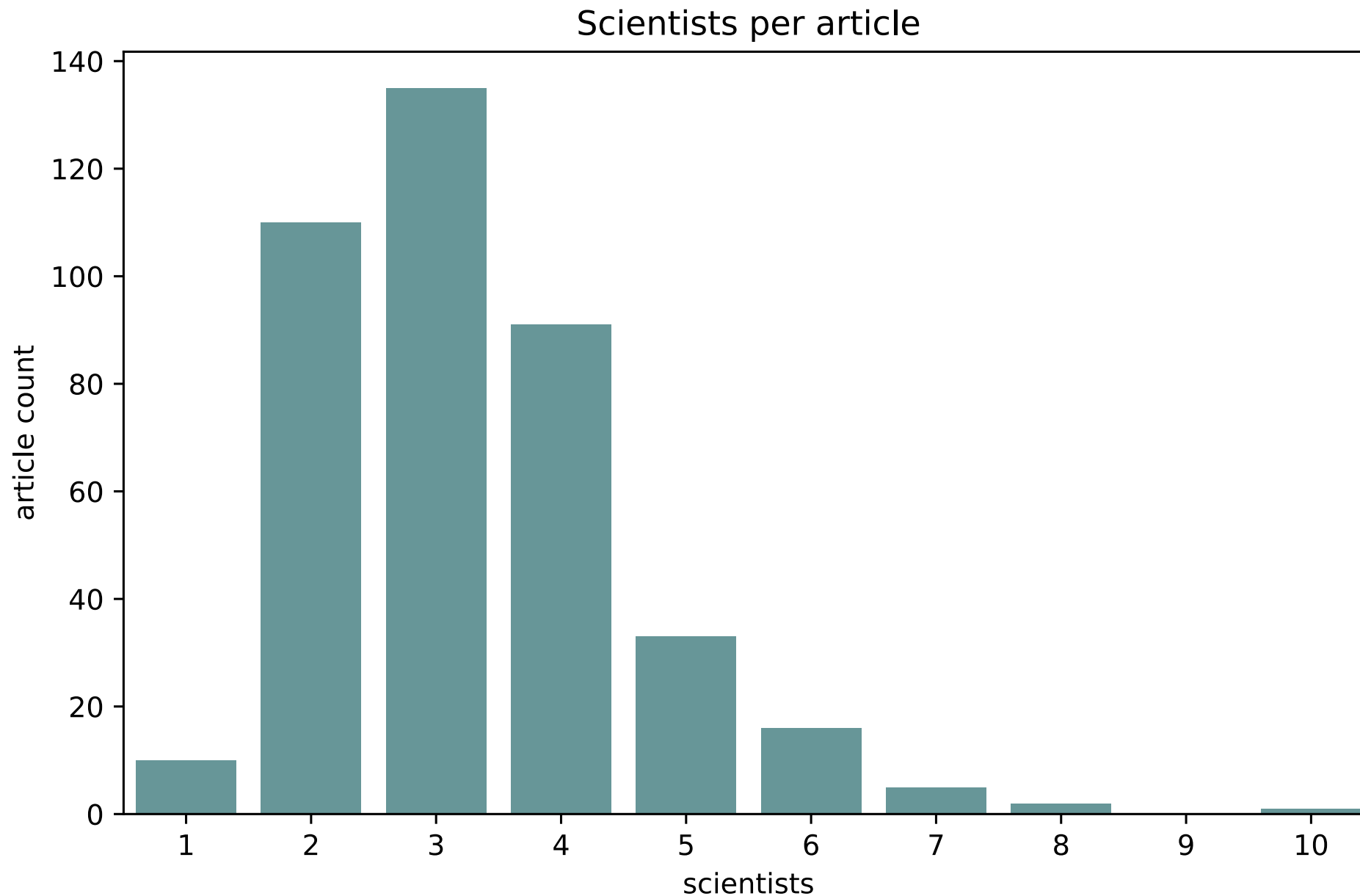
Некоторые очень стараются...



Авторы, которые участвовали в наибольшем количестве статей

# Массовость статей

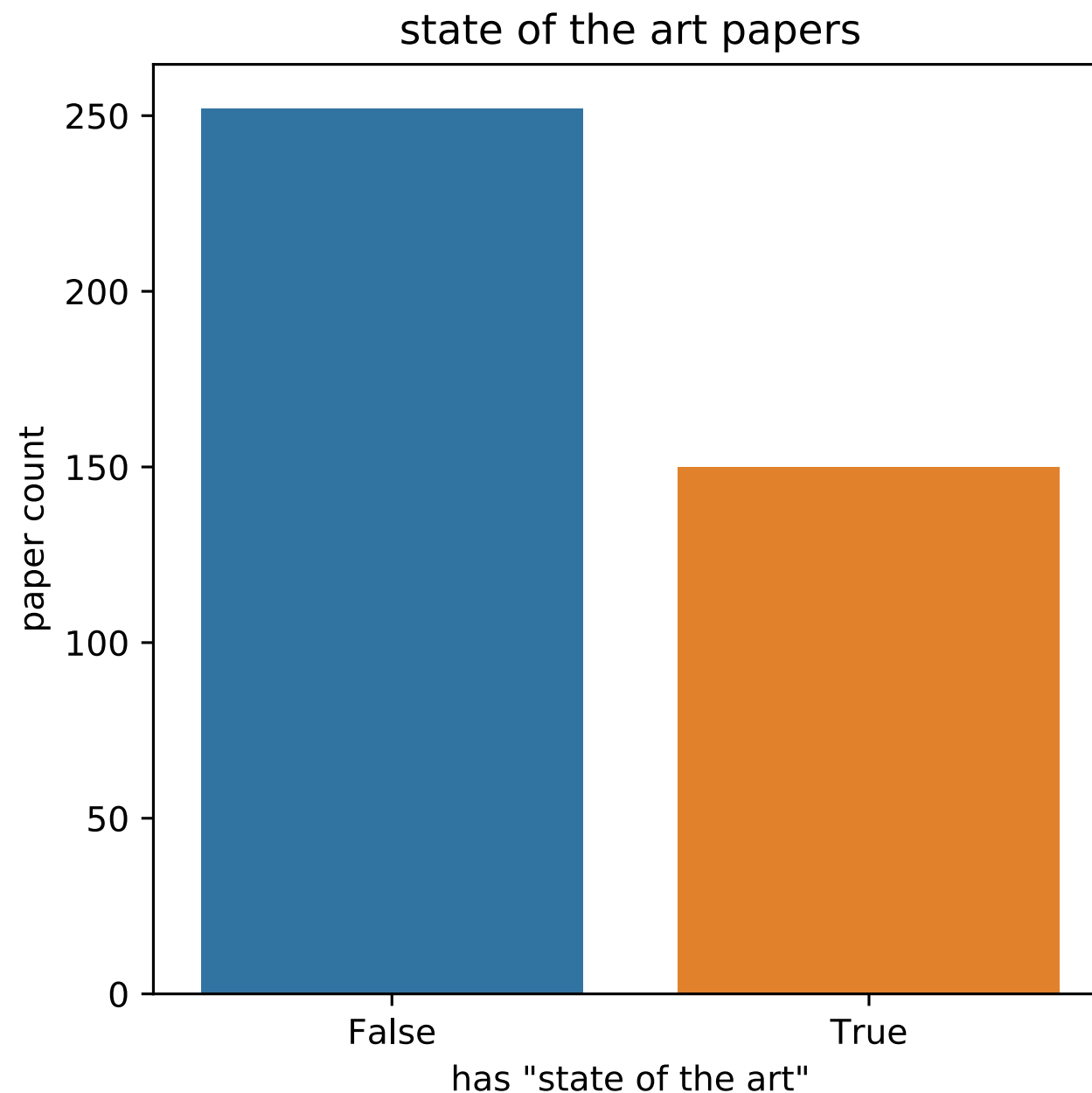
Ничего особенного — обычно 2-4 человека пишут одну статью



Сколько авторов у одной статьи

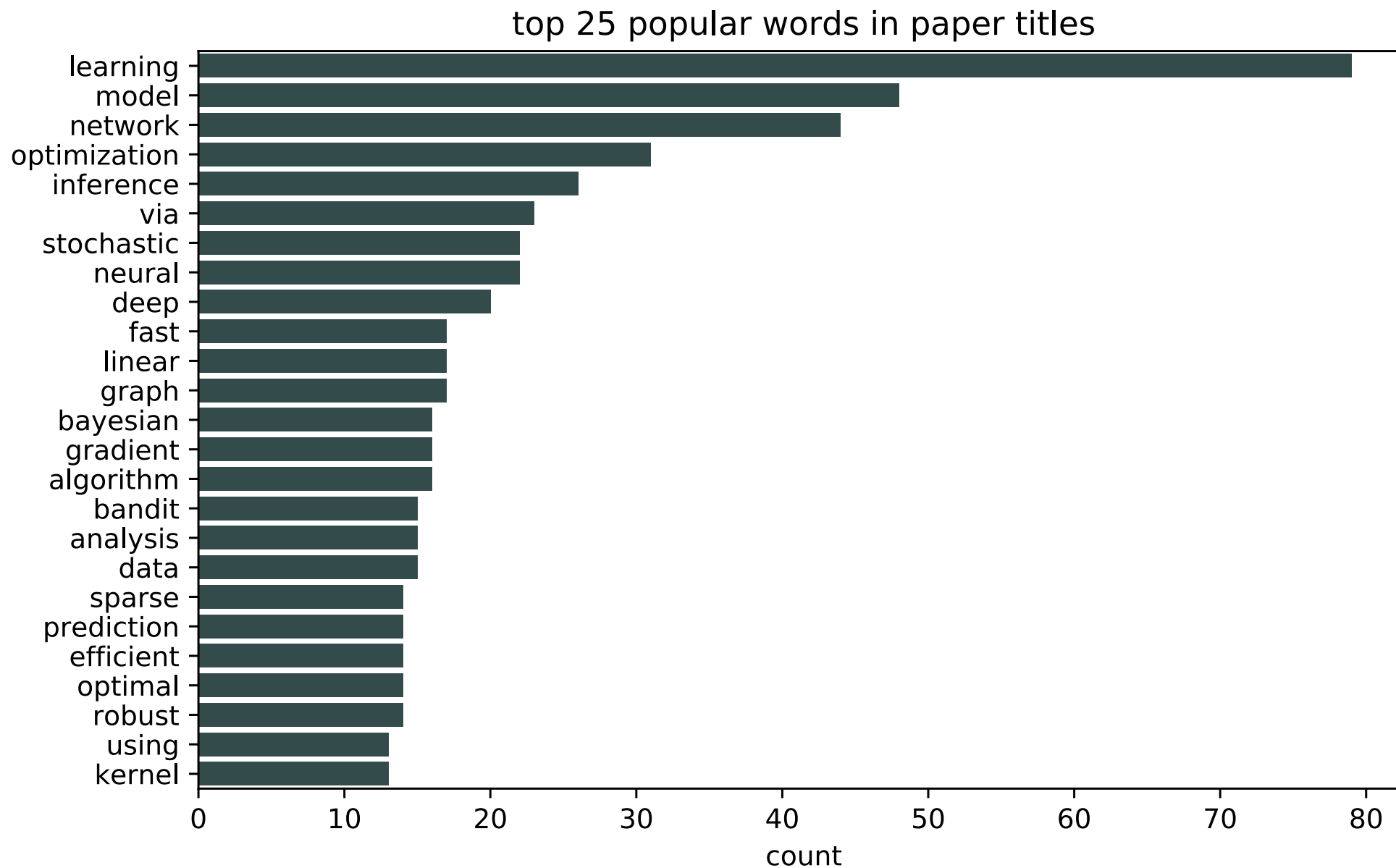
# State-of-the-art статьи

Много «революционного»



Количество статей, упоминающих «state-of-the-art»

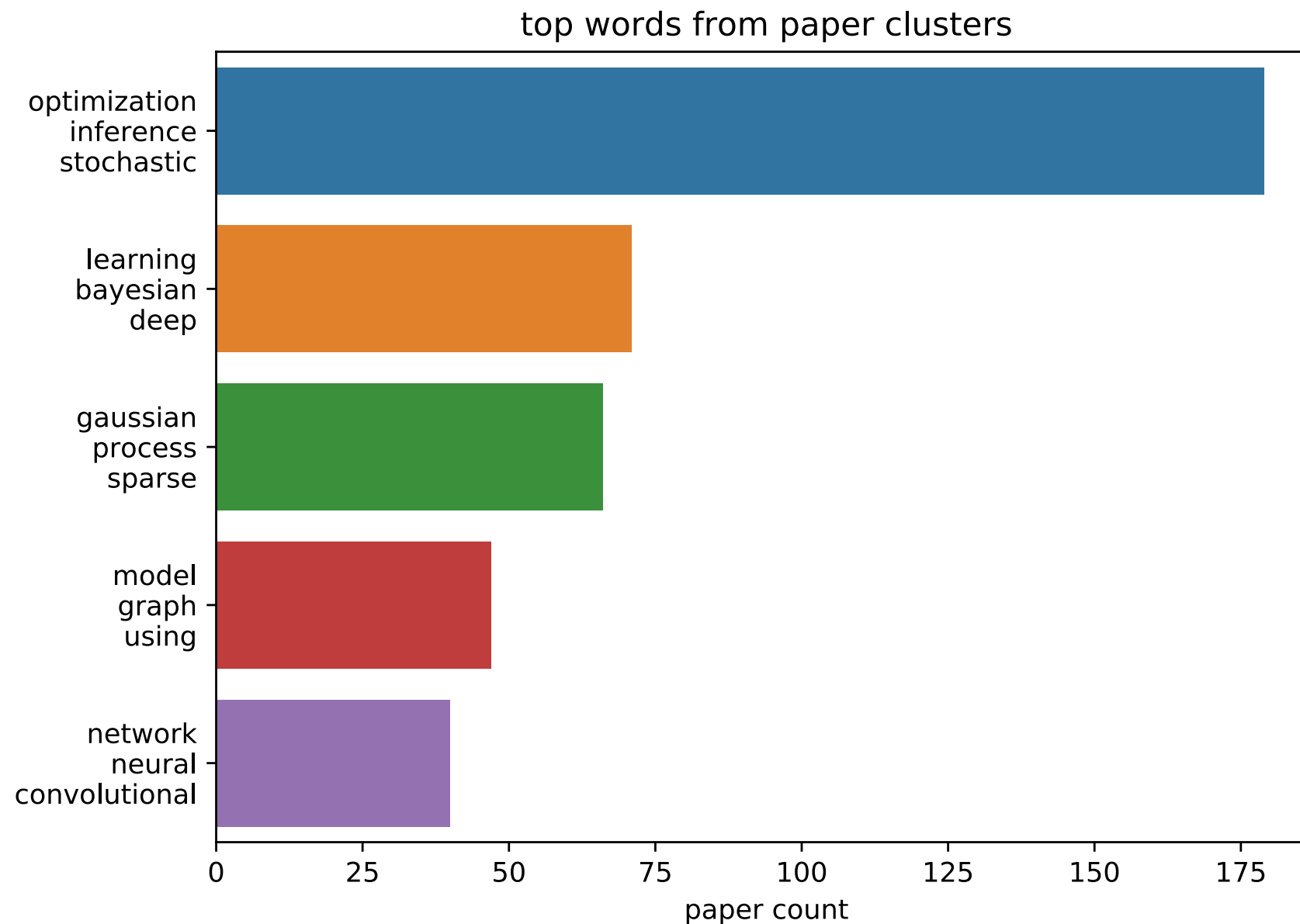
# Наиболее популярные слова в заголовках



Топ популярных слов в заголовках без стоп-слов и после лемматизации

# Характерные статьи

Тренды исследований в 2015



Три наиболее популярных слова среди пяти кластеров статей, образованных с помощью kmeans на tf-idf преобразованных заголовках