

Analysis of Reddit Posts

Alireza Karimi



Project Objectives

- Collecting posts from two subreddits
- Utilizing NLP to train a classifier in order to predict the subreddit of a given post

Science



Psychology



Scraping

- Using Pushshift API
- Collected 10,000 posts in Science subreddit and 9,991 posts in Psychology subreddit
- Stored the Id, Title, Text, and Subreddit of each post
- Concatenated the title and the text of each post
- Each subreddit word bank contained more than 1 million words.



Science Subreddit

The word cloud illustrates the following themes:

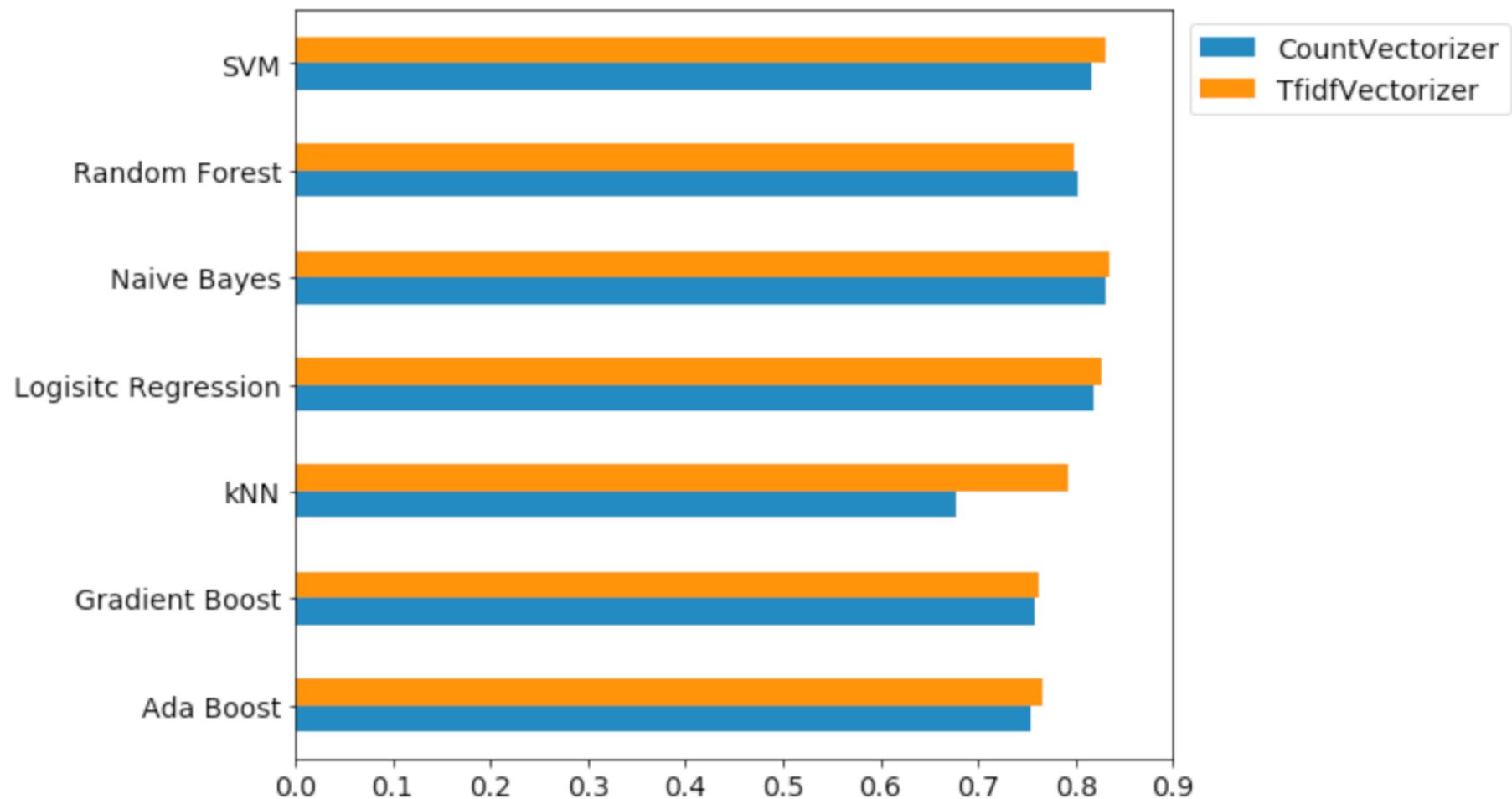
- Health and Medicine:** researcher, new, children, drug, study, disease, genetic, cancer, health.
- Technology and Science:** human, first, world, protein, change, find, test, suggest, reveal, animal, increased, common, ancient, bacteria, best, gene, plant, control, planet, structure, known, nasa, high, area, condition, material, species, nasa, high, area, much.
- Society and Behavior:** according, linked, create, climate, population, type, better, years ago, analysis, social, result, rate, sleep, help, around, body, ocean, space, two, become, suggests, new, percent, men, part, discovered, find, work, age, even, including, three, increase, increase, still, year, system.
- Environment and Ecology:** new, study, find, compared, mars, large, early, many, global, reveal, among, state, study, suggest, come, take, well, enough, common, ancient, behavior, current, reduce, earth, dna, individual, cost, might, evidence, region, say, chemical, half, information, without, effective.
- Psychology and Neuroscience:** scientist, people, brain, science, time, report, good, development, mice, university, climate, change, cause, now, level, studies, finds, new, star, cell, thought, key, first time, much.

Psychology Subreddit



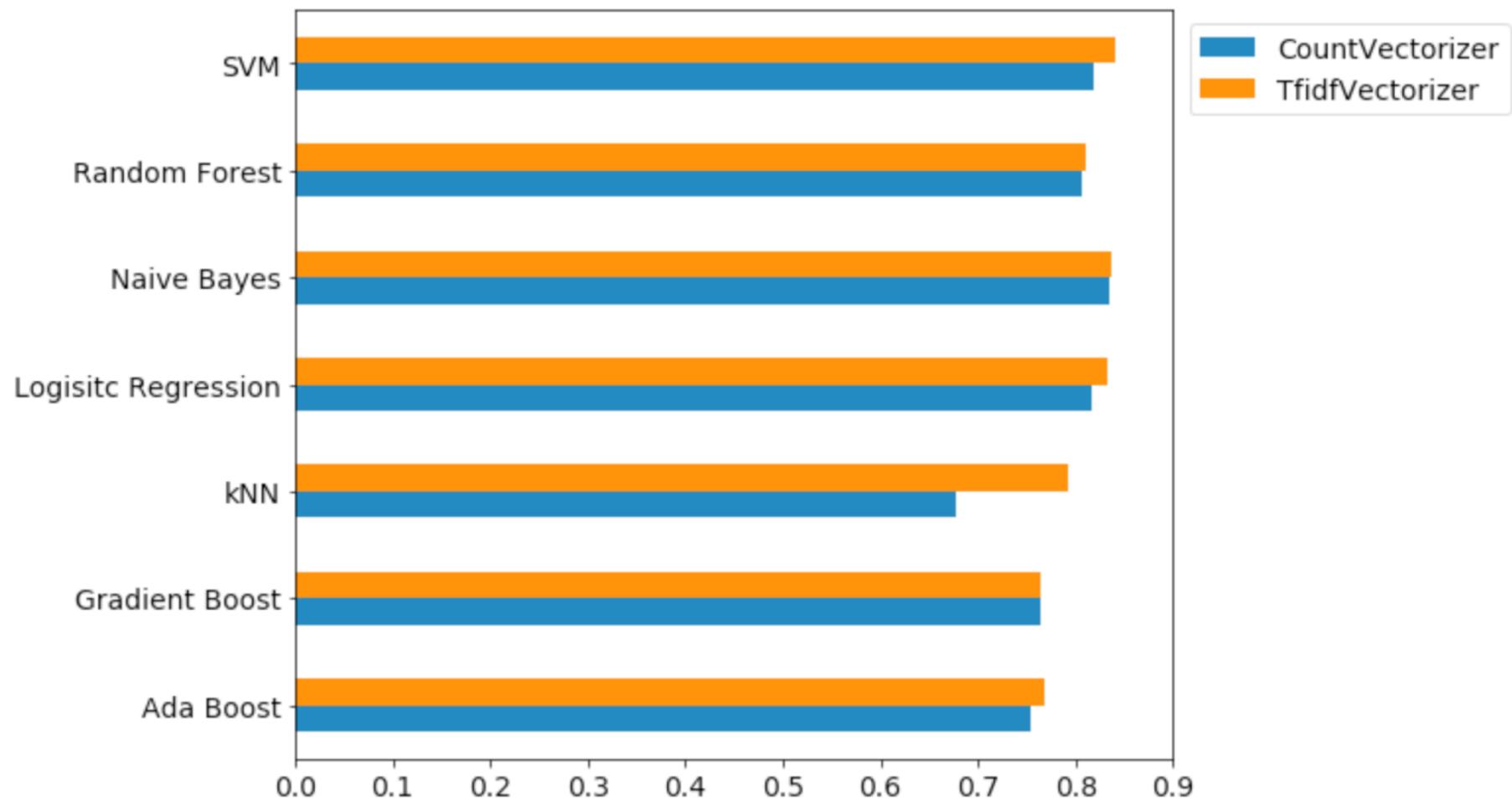
Modeling

Accuracy of the Train set



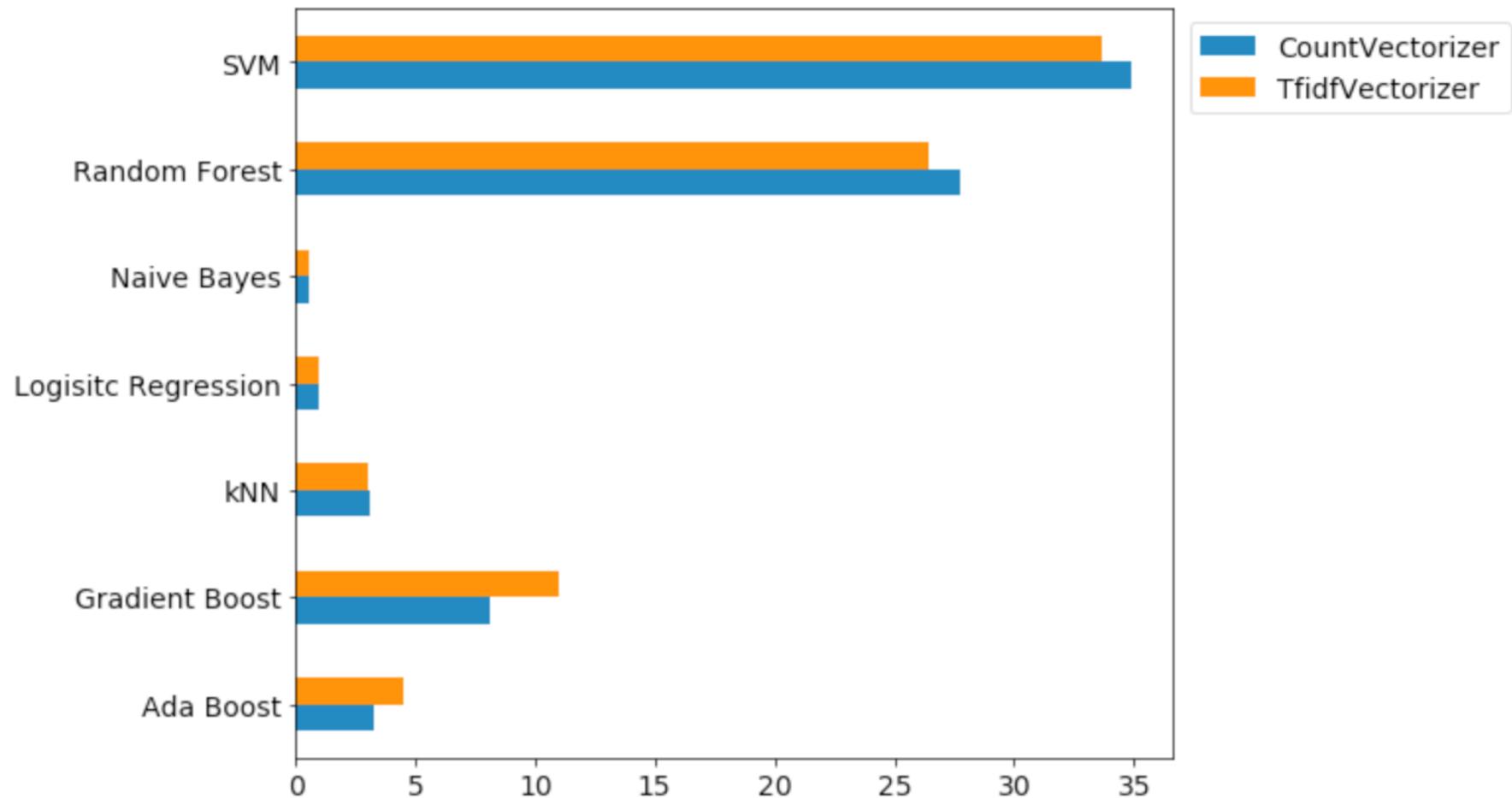
Modeling

Accuracy of the Test set



Run Time of Models

Run Time per Fit (s)



False Predictions

 Join the discussion [BECOME A REDDITOR](#) X

Posted by u/mvea MD-PhD-MBA | Clinical Professor/Medicine 1 month ago

2.0k The link between cannabis use and psychological distress is stronger for women than men, according to new research in The Canadian Journal of Psychiatry (Canada recently legalized recreational cannabis). The findings indicate that negative effects of cannabis use may be more pronounced for women.

psypost.org/2018/1...

[Psychology](#)



COMMUNITY DETAILS

 r/science
20.2m Subscribers 7.5k Online

This community is a place to share and discuss new scientific research. Read about the latest advances in astronomy, biology, medicine, physics, social science, and more. Find and submit new publications and

 Join the discussion [BECOME A REDDITOR](#) X

Posted by u/MightyViral 2 months ago

1 Which Candy Best Represents Your Personality?

mightyviral.com/which-...



COMMUNITY DETAILS

 r/science
20.2m Subscribers 7.2k Online

This community is a place to share and discuss new scientific research. Read about

 Join the discussion [BECOME A REDDITOR](#) X

Posted by u/cannabiscannabis999 3 months ago

1 5 Steps to overcome shyness with women

goodheed.com/overco...

[Social Science](#)



COMMUNITY DETAILS

 r/science
20.2m Subscribers 7.6k Online

This community is a place to share and discuss new scientific research. Read about