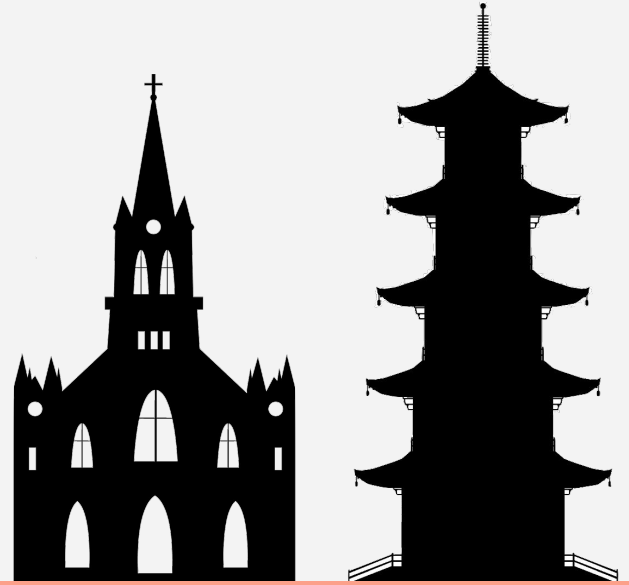


LINGUISTIC PATTERN ANALYSIS ON SUBREDDIT POSTS



On r/Buddhism and r/Christianity

by : Jefferson Qiu

01

INTRODUCTION

Analysis on linguistic pattern on subreddit

02

ON THE SURFACE

What is the macro differences based on EDA?

03

DIVING DEEPER

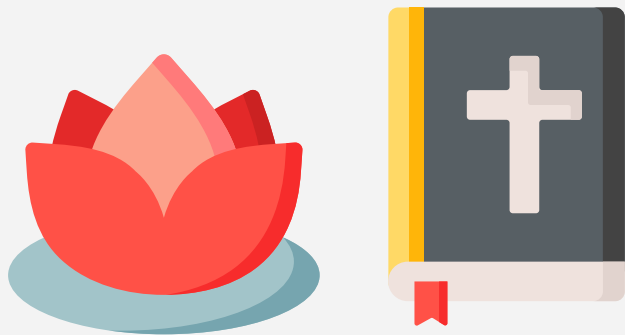
How the language usage is different?

04

THE SUMMARY

So what?

01



INTRODUCTION

BACKGROUND

- **Subreddit** is a niche communities to discuss certain topics
 - Has unique linguistic pattern and culture
 - Conducive for the fanbase, foreign for new users
- **Problem Statement**
 - Can analysis through EDA and NLP uncover subtle culture in subreddits?
- **Case on Point:**
 - [reddit.com/r/Buddhism](https://www.reddit.com/r/Buddhism) and [reddit.com/r/Christianity](https://www.reddit.com/r/Christianity)
 - Religion-based subreddit
 - Different underlying doctrine

THE APPROACH



WEB-SCRAPING

Scrape title, post and comments



DATA SELECTION

Where to find linguistic pattern?



EDA

Any macro pattern on the text?



APPLICATION OF NLP

Model selection and comparison



INSIGHT FROM NLP

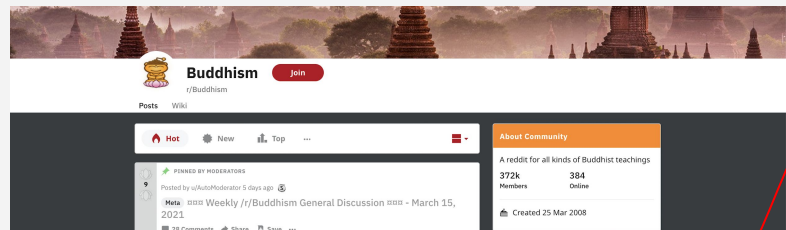
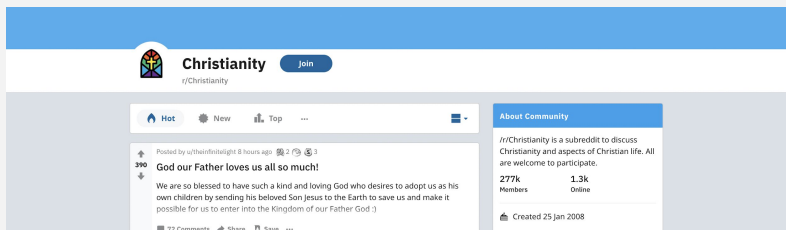
Identification of jargons from each subreddit



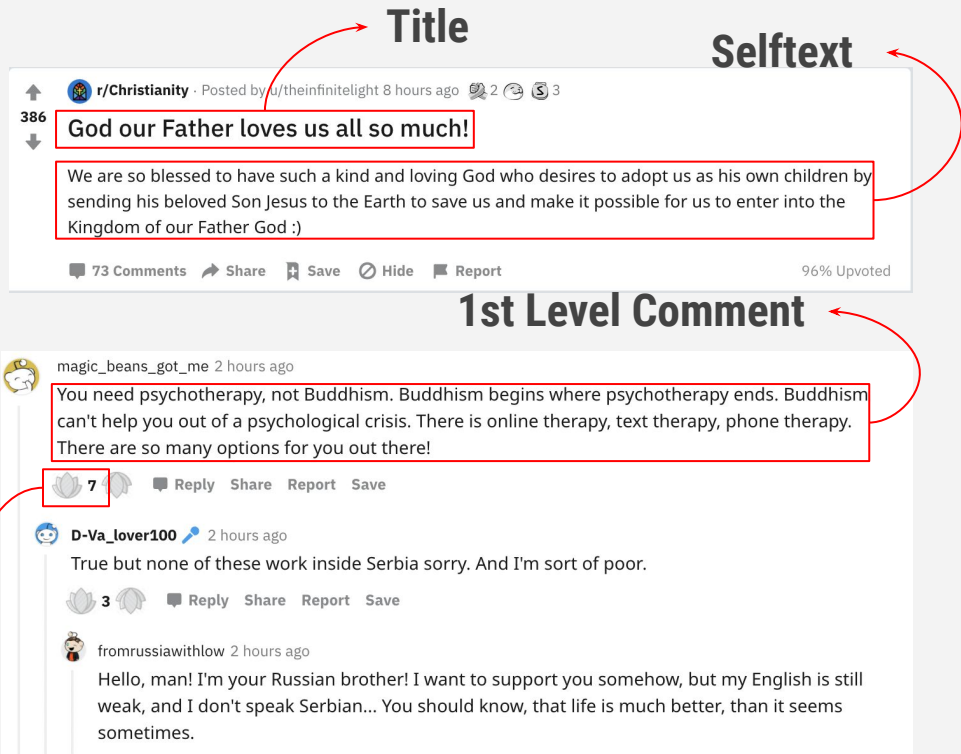
MISCLASSIFIED POSTS

Any loophole in the model/analysis

WEB-SCRAPING THE SUBREDDITS

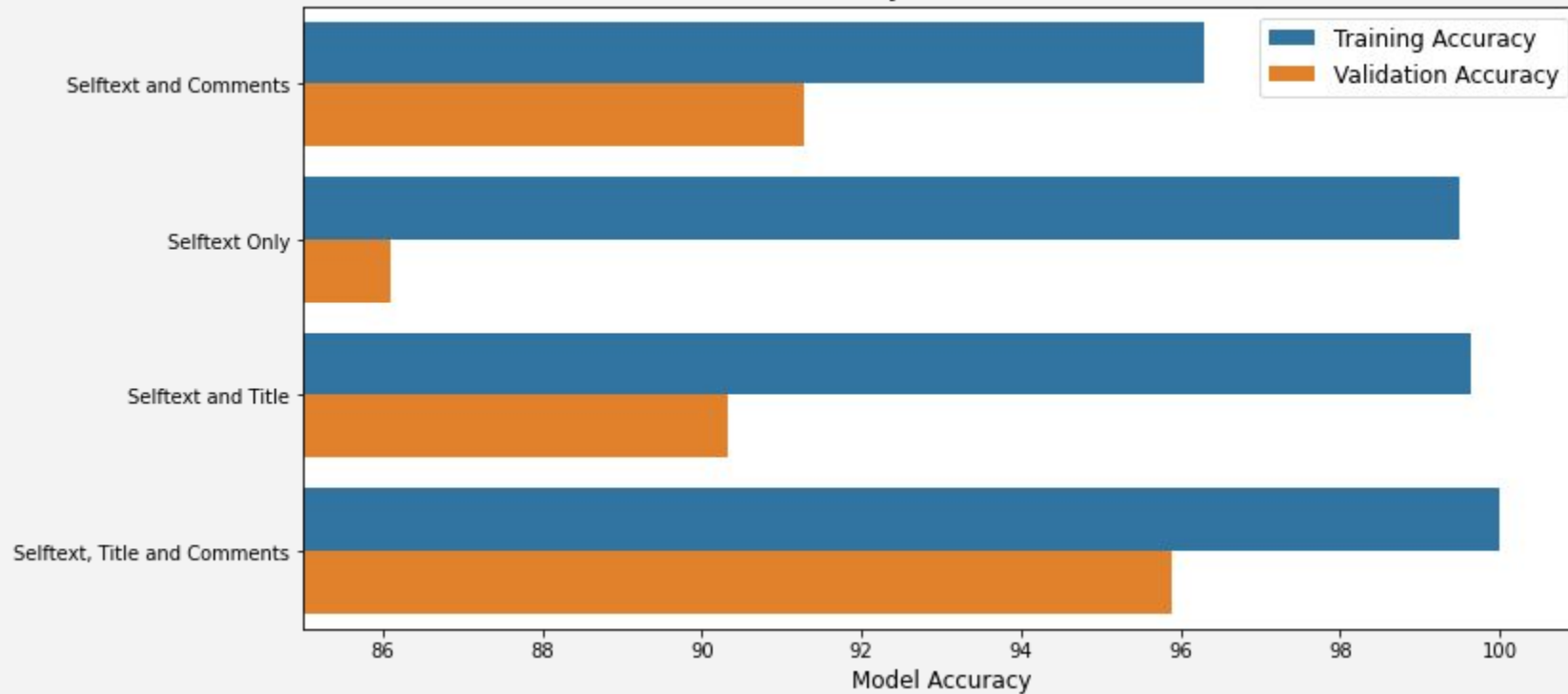


Upvote!

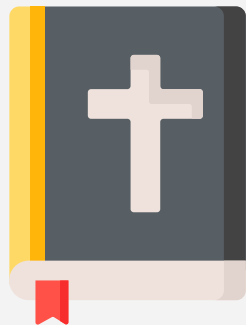


DATA SELECTION

Model Accuracy for Various Datasets

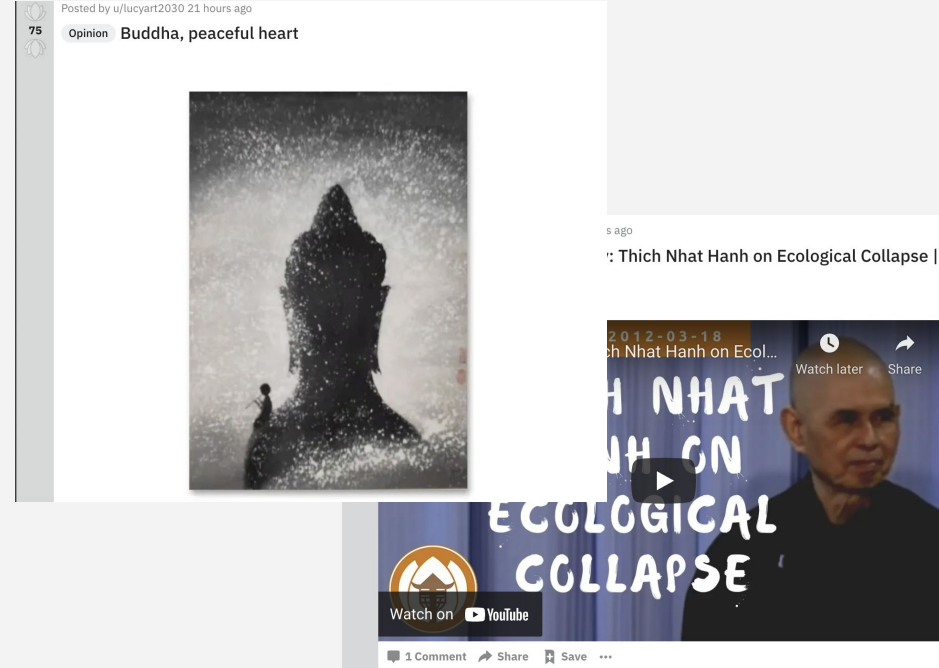
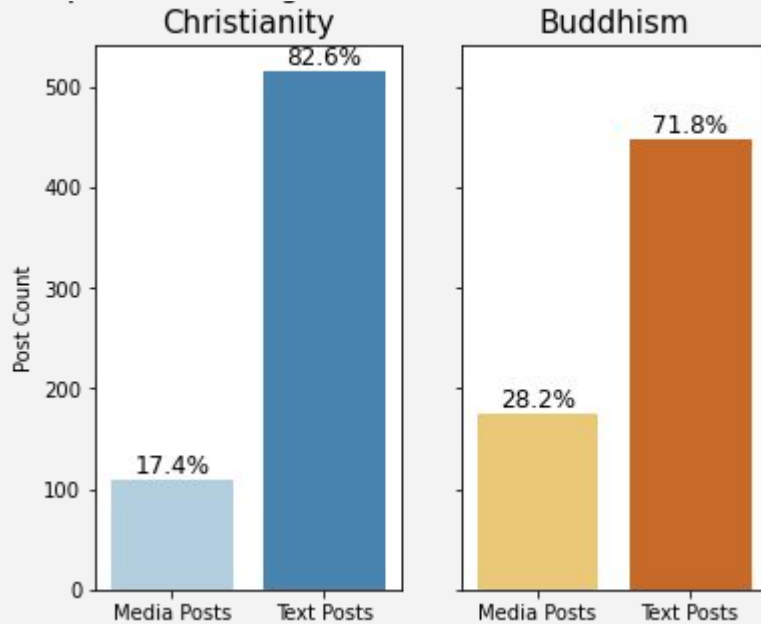


02



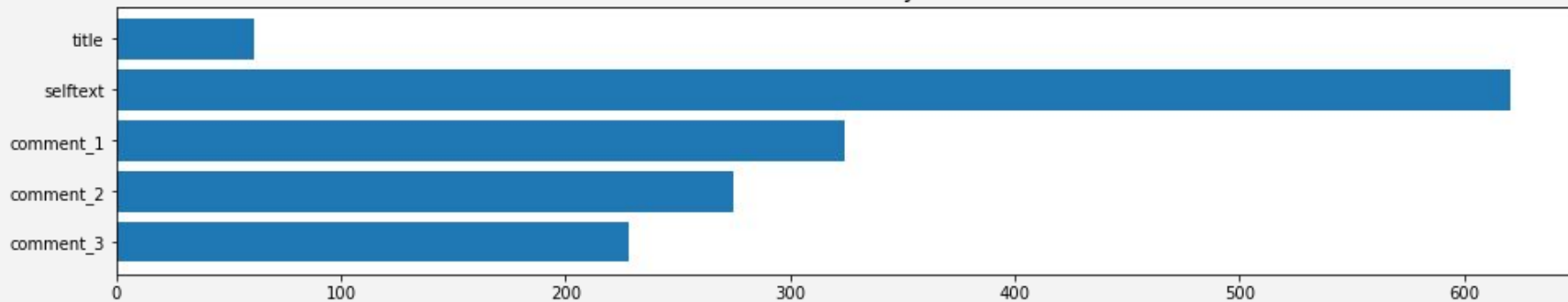
ON THE SURFACE

EDA: EMPTY POSTS

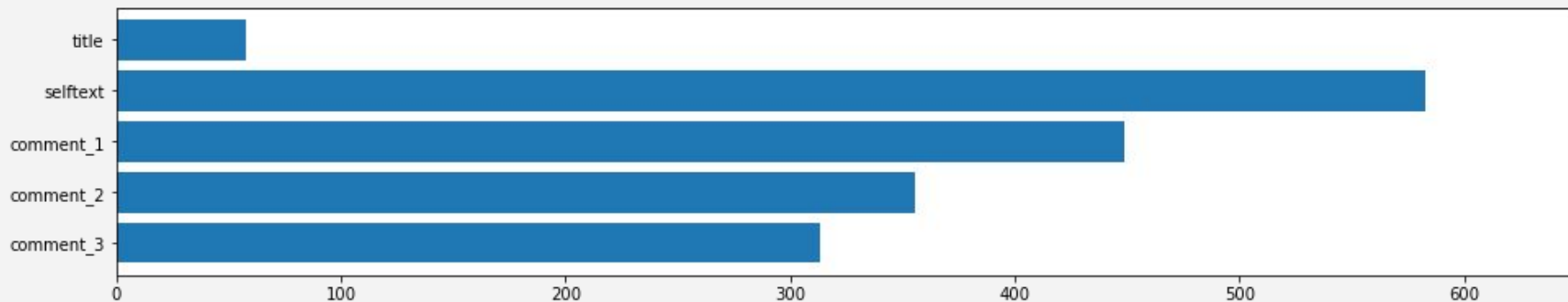


EDA: POST LENGTH

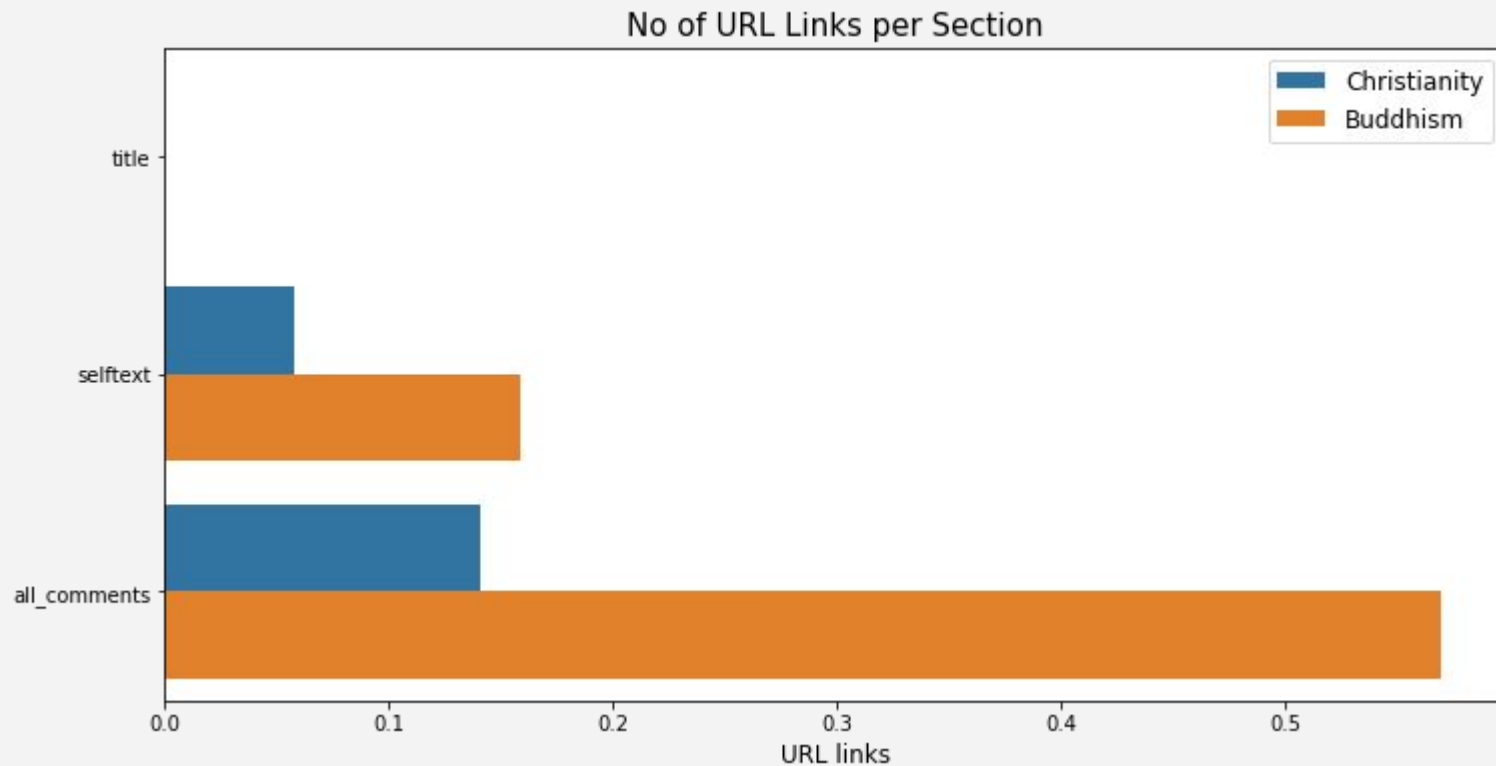
Christianity



Buddhism



EDA: URL SHARING



EDA: FREQUENTLY-USED WORDS



*HOW GREAT IS OUR **GOD**!!!! ... For most of my life I didn't have a close relationship with **God** until this year I felt a strong awakening in myself. At the time I was devastated ..., but I realized this sudden conviction was a result of my absent relationship with **God**, ...*

*I want to be close with **God** but I feel that my porn addiction has really badly affected my mind and plus somehow I get intrusive thoughts insulting **God** and the bible as well as other things! I don't want to think that I don't agree with those thoughts insulting **God** ... make me feel like I've f'ed up my chances of being with **God** ...*

EDA: FREQUENTLY-USED WORDS



One can engage with the Dharma as **one** would like with any view whatsoever, ... Of those, right view is the forerunner. And how is right view the forerunner? **One** discerns wrong view as wrong view, and right view as right view. This is **one's** right view. And what is wrong view? 'There is nothing given, nothing offered, nothing sacrificed.

... There is no reason why mindfulness should be different from focusing all **one's** attention on **one's** work, to be alert and to be using **one's** best judgment. During the moment **one** is consulting, resolving, and dealing with whatever arises, a calm heart and self- control are necessary if **one** is to obtain good results.

03



DIVING DEEPER

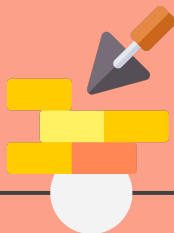
What NLP methods reveals?

NLP AS A MEANS



STEP 01

Combine posts from both subreddits



STEP 02

Build **NLP models to classify** posts to their original subreddits



STEP 03

Select best **NLP with criteria**:

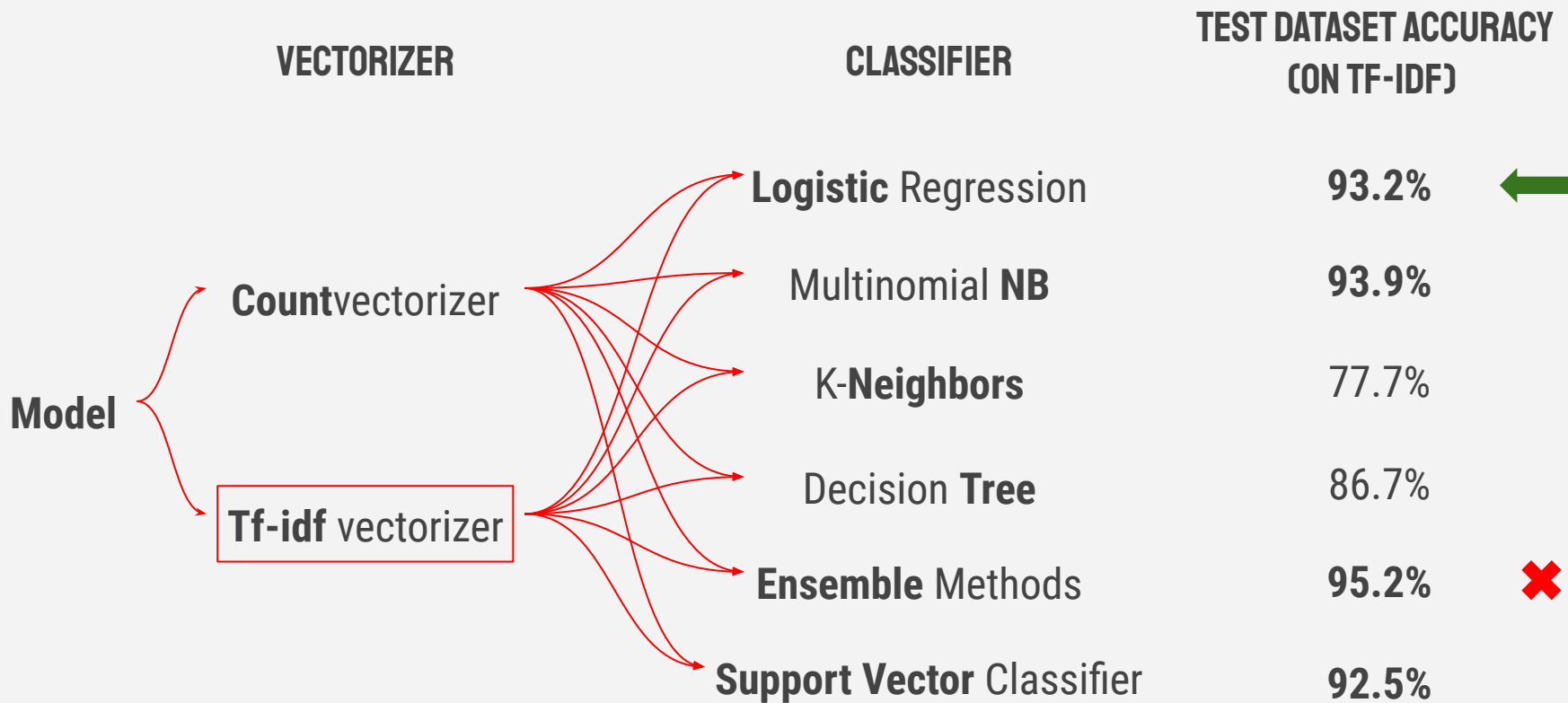
- High **interpretability**
- High **accuracy**



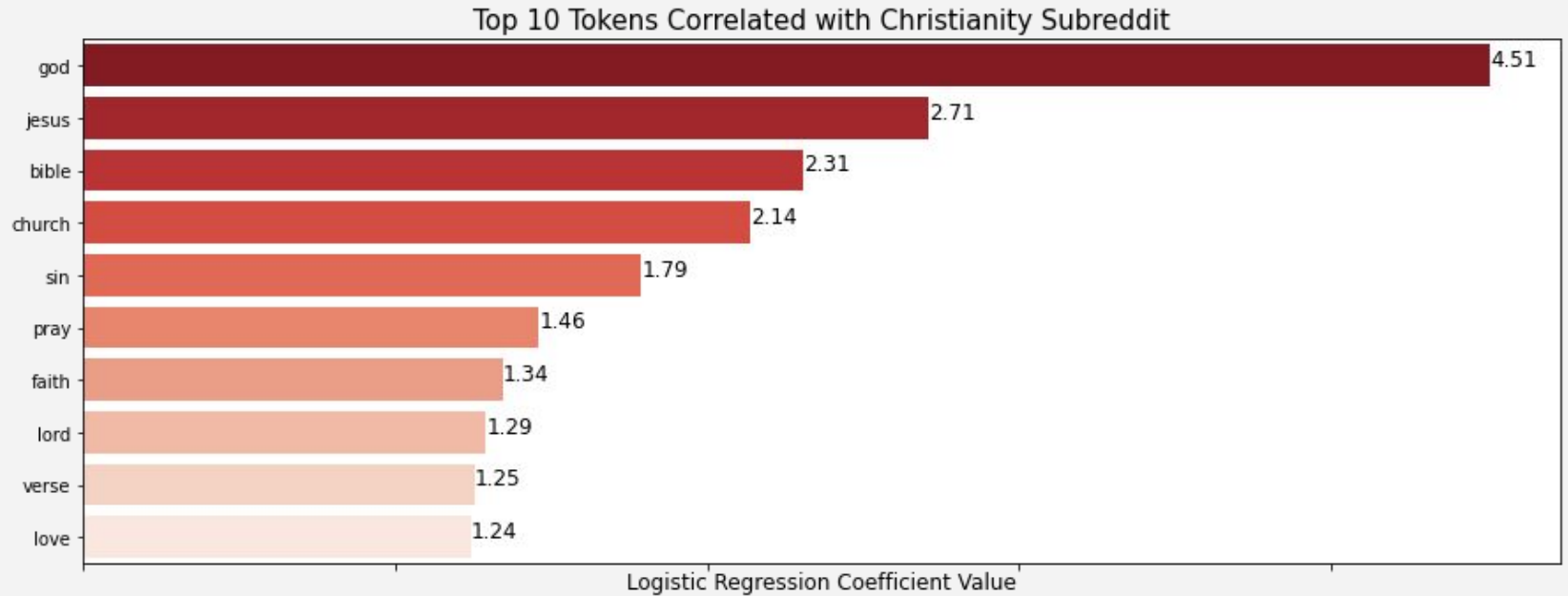
STEP 04

Deduce **jargons** used correlated to each subreddits

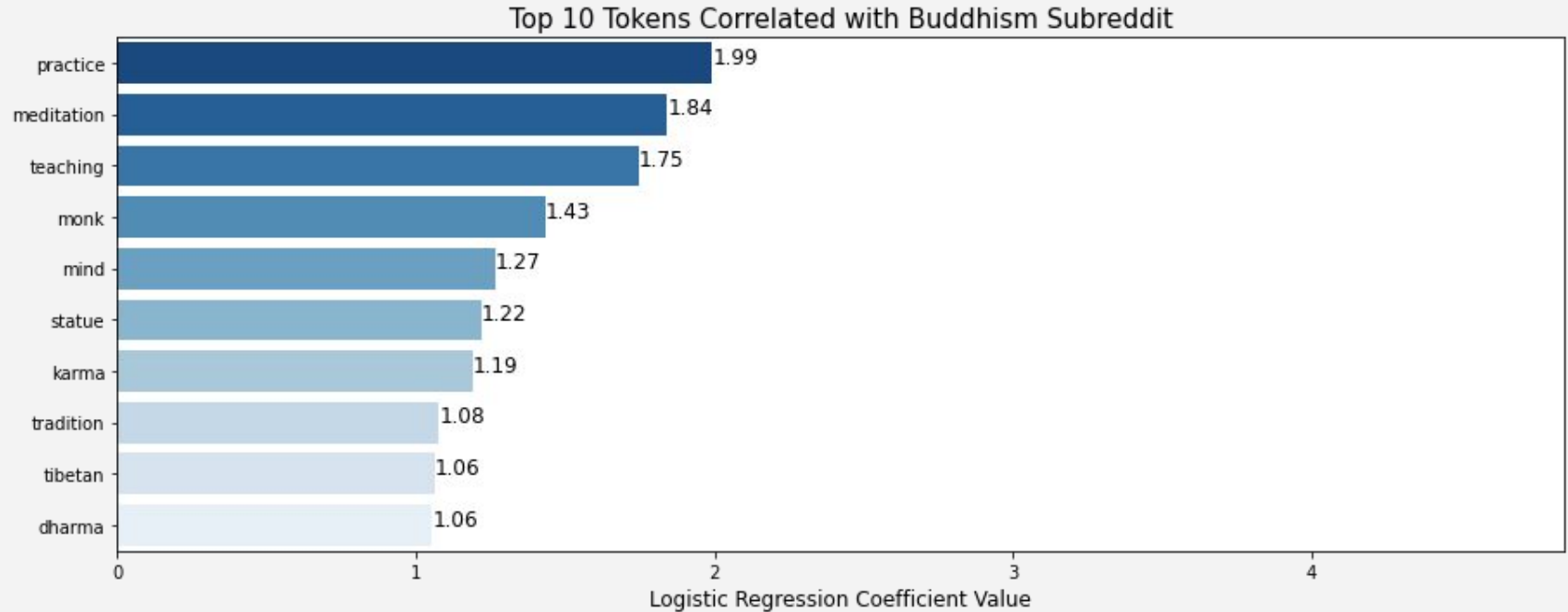
BUILDING NLP MODEL



IDENTIFICATION OF JARGON



IDENTIFICATION OF JARGON



IDENTIFICATION OF JARGON

Topic	Token Primarily used in Christian Subreddit	Token Primarily used in Buddhism Subreddit
Teacher/ Supreme figure	Lord, God	Buddha
Knowledge Transmission and Repository	Bible	Teaching, Dharma
Object/ Place	Verse, Church	Statue, Tibetan
Doctrine	Faith, Love, sin	Karma, Mind
Activity	Pray	Practice, Meditation

SENTIMENT ANALYSIS

	Christianity SIA	Buddhism SIA
title	0.058067	0.079068
selftext	0.198783	0.234400
all_comments	0.414328	0.433278

- Original Post:

*Last year in August, I had a huge crush on a girl that **rejected** me because of my **creepy/needy** behavior. I've also told her best friend some **embarrassing** things I've done as a child. The crush turned into an **obsession** without me knowing it....*

- Comments:

*Life is a process without immediate gratification ... identify the issue and make slow and **steady** steps towards change. .. you **HAVE** to have **compassion** on yourself as if you are another person, as if you were someone else you ***love** and have **compassion** for .*

MISCLASSIFIED POSTS

True Label	Christianity	Buddhism
	Christianity	Buddhism
Christianity	192	15
Buddhism	9	197

MISCLASSIFIED POSTS

Buddhism Subreddit Post **Classified** as from **Christianity** Subreddit

Original Post:

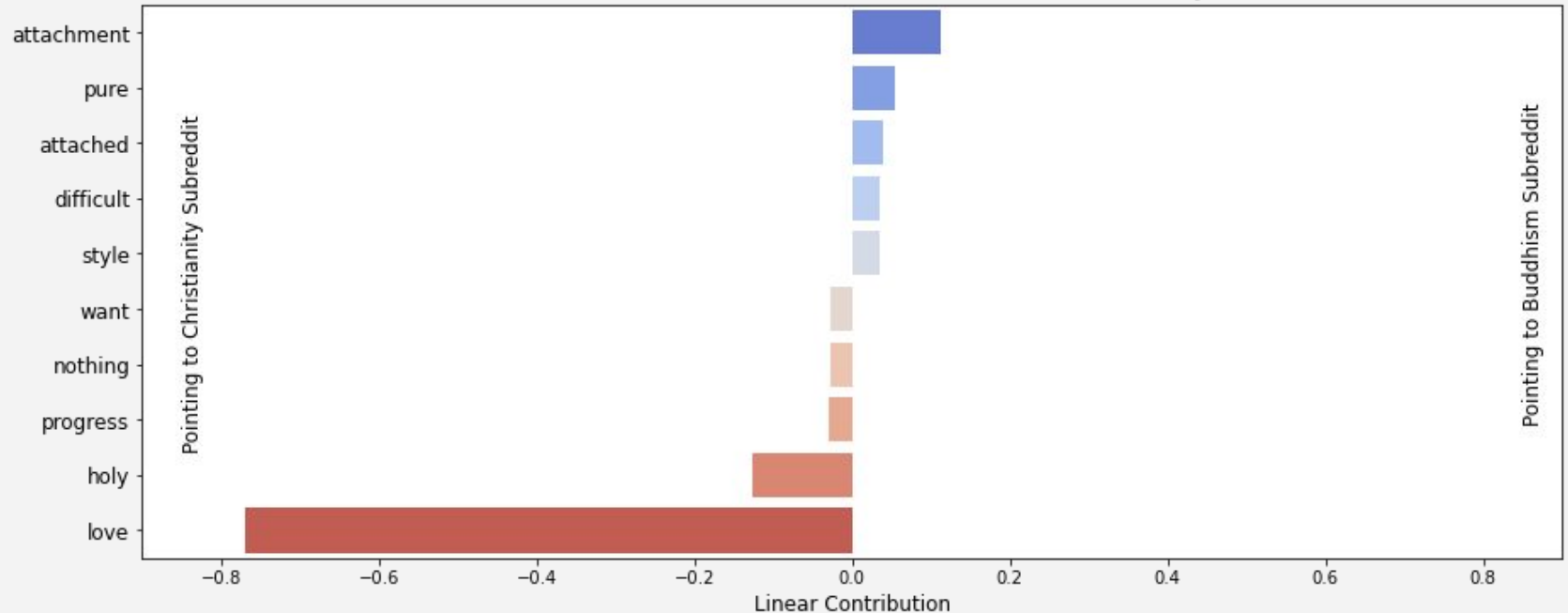
*My question will be straightforward: how do you reconcile having goals of spiritual realization with the romantic love that you can have when you are not a **monk**?*

*If it is possible. Maybe it's not possible, and maybe it's just my ego, with all its attachments, looking for a way to hold on. Any attachment to the world is illusory, and temporary, and can therefore only lead sooner or later to suffering and/or dissatisfaction (**dukkha**). ...*

*I often receive the answer that, up to a certain point, the two are reconcilable, but this answer has always left me hungry: if, ultimately, romantic **love** is an **attachment** and we have to be **attached** to nothing to be **holy**, it is difficult to **love** fully knowing this. It is just as difficult to **love** fully when you know that personalities are illusory and mortal (**anatta**)....*

MISCLASSIFIED POSTS

Buddhism Subreddit Post **Classified** as from **Christianity** Subreddit



MISCLASSIFIED POSTS

Christianity Subreddit Post **Classified** as from **Buddhism** Subreddit

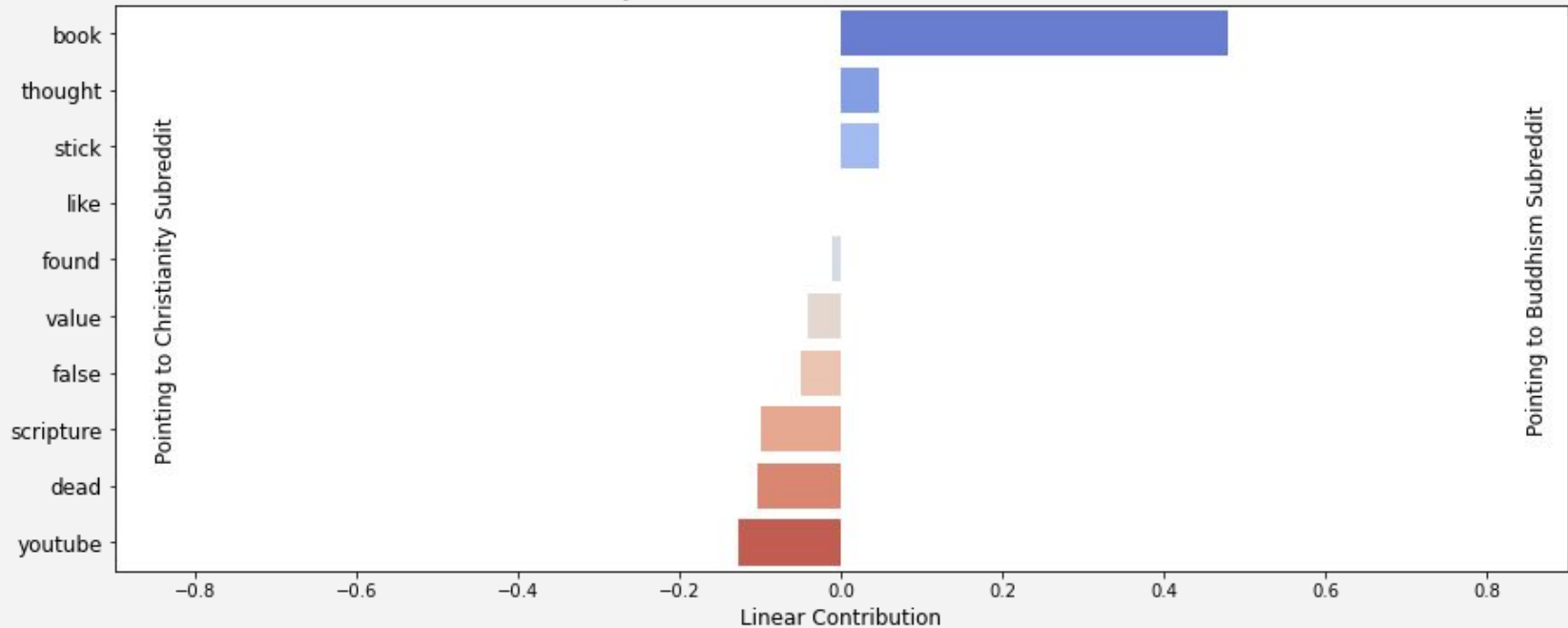
Original Post:

I'd like to get some of you guy's thoughts on the book of Enoch. I believe it was discovered around the 1700s then again when the Dead Sea scrolls were found. How should this book be viewed? Is it false teaching? What value does this book bring? Thoughts on the book of Enoch.

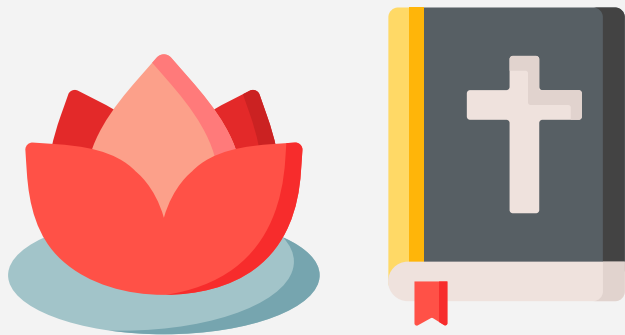
<https://www.youtube.com/watch?v=440hI9lnAlc> I think we should stick to scripture. Enoch, and books of apocrypha are not in it

MISCLASSIFIED POSTS

Christianity Subreddit Post **Classified** as from **Buddhism** Subreddit



04



SUMMARY

THE SUMMARY

- People on **Buddhism** subreddit shares more **media** posts, i.e. image, videos, links.
- People on **Christianity** subreddit shares a longer post
 - But people on **Buddhism** subreddit replies in longer comments
- **Jargons** are specific to the religion, but **topic** are similar
 - **Topics**: religion founder, knowledge repository, believer, doctrine, and activity
- Both subreddits has a **positive sentiment analysis**
 - Especially on **comments** section as subreddits console each other

MODEL LIMITATION AND RECOMMENDATION

- Misclassified when no mention of **jargons**
 - Especially if jargons of other class are used
- Model **missing on context** as word count is used
- **Recommendations:**
 - To deploy human interpreter on selected posts
 - To deploy context dependent NLP
 - BERT, ELMo

THE CONCLUSION

- On the **surface** there are many **difference** on both subreddits
 - Length of post
 - Content (media or text or external links)
 - The jargons used
- On the **deeper** level, there are many similarities also!
 - Tend to be positive note
 - Jargons are talking about similar topics
- EDA and NLP is able to **uncover many unspoken culture** on both subreddits

THANKS!

Do you have any questions?

