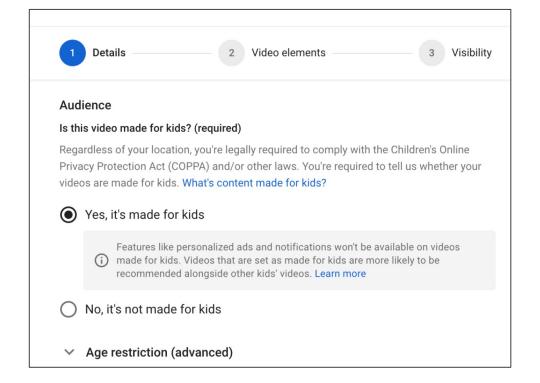
Identifying YouTube Videos "Made for Kids"

Sue Lim July 21, 2023

I. PROJECT BACKGROUND

YouTube Videos "Made for Kids"

- FTC requires identifying YouTube videos "made for kids" under COPPA (Children's Online Privacy Protection Act)
- Such videos are under certain restrictions
- YouTube may face legal consequences if content is not categorized correctly
 - **→** The need to flag videos correctly



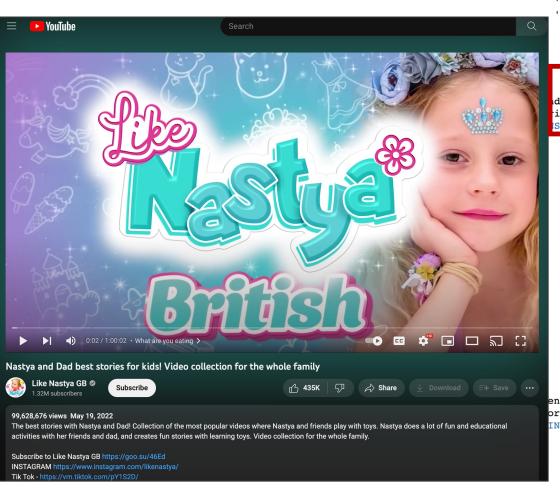
Recommendation

Use a Naïve Bayes model

1. It has high prediction accuracy

2. The model can save manual human efforts in identifying videos "made for kids"

YouTube API Data



```
{ 'kind': 'youtube#videoListResponse',
 'etag': '8EdUuit1u9GHpB-mRuylJLQQqEk',
 'items': [{'kind': 'youtube#video',
   'etag': '1EXzyGvBtQsYeIhPTWILufbNAxE',
   'id': 'UQyHkp8-B6E',
   'snippet': { 'publishedAt': '2022-05-19T09:00:11Z',
    'title': 'Nastya and Dad best stories for kids! Video collection for the whole family',
    <mark>'description':</mark> 'The best stories with Nastya and Dad! Collection of the most popular videos where Nastya and frie
 ds play with toys. Nastya does a lot of fun and educational activities with her friends and dad, and creates fun sto
 ies with learning toys. Video collection for the whole family.\n\nSubscribe to Like Nastya GB https://goo.su/46Ed\nI
 STAGRAM https://www.instagram.com/likenastya/\u200b\nTik Tok - https://vm.tiktok.com/pY1S2D/',
     tnumbnalls : { derault : { url : https://l.ytlmg.com/v1/UQyHKP8-B6E/derault.]pg ,
    'channelTitle': 'Like Nastya GB',
    'tags': ['Like Nastya GB',
     'Nastya',
     'Like Nastya',
     'Funny stories for kids',
     'Kids videos',
     'Nastya and Dad',
     'Kids channel',
     'Entertainment for kids',
     'best stories for kids',
     "children's educational videos"
     'Fun with friends',
     'Kids Video Collection',
     'educational toys',
     'Pretend to play',
     'Kids pretend to play'],
    categoryia: 22,
    'liveBroadcastContent': 'none',
    'localized': {'title': 'Nastya and Dad best stories for kids! Video collection for the whole family',
     'description': 'The best stories with Nastya and Dad! Collection of the most popular videos where Nastya and fri
ends play with toys. Nastya does a lot of fun and educational activities with her friends and dad, and creates fun st
ories with learning toys. Video collection for the whole family.\n\nSubscribe to Like Nastya GB https://goo.su/46Ed\n
INSTAGRAM https://www.instagram.com/likenastya/\u200b\nTik Tok - https://vm.tiktok.com/pY1S2D/'}},
   'status': {'uploadStatus': 'processed',
    'privacyStatus': 'public',
    'license': 'youtube',
    'embeddable': True,
    'publicStatsViewable': True,
    madeForKids : True}}],
 pageinfo : { totalkesults : 1, resultsPerPage': 1}}
```

Benefits of Machine Learning

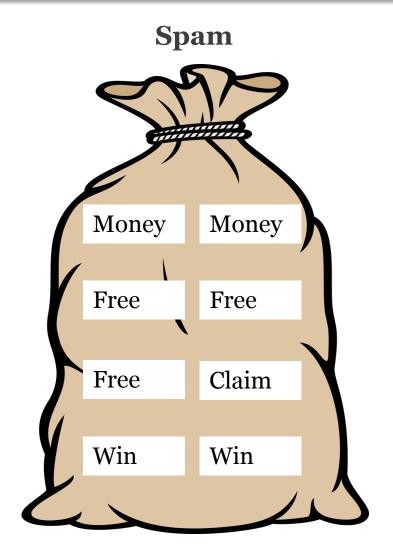
Highly scalable

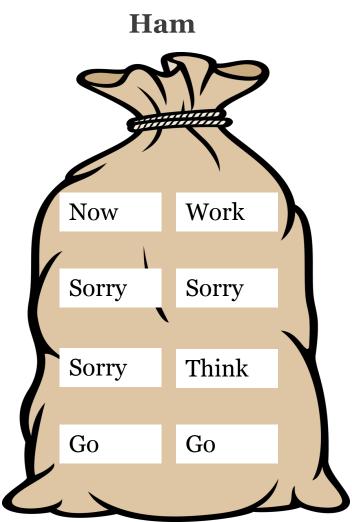
Reduce manual human efforts

Consistent criteria in identifying videos

II. MAIN MODEL

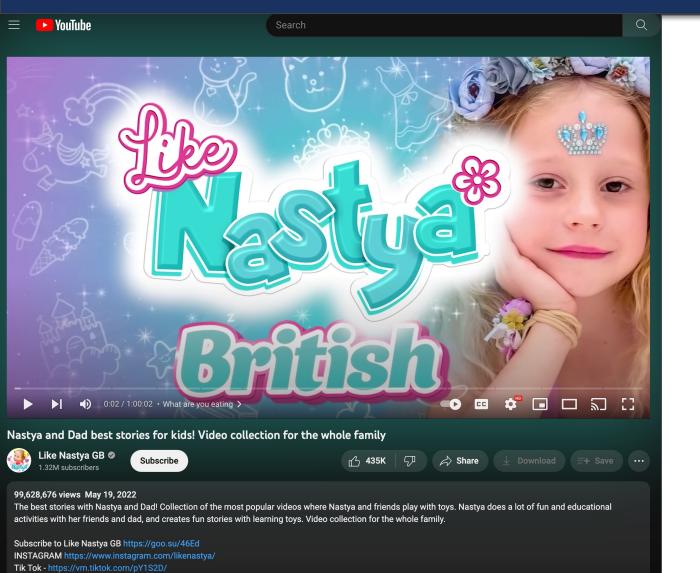
Naïve Bayes Classifier





Then, is "Win \$30" Spam or Ham?

Naïve Bayes Classifier – Nastya Example



Keywords from Nastya Example: nastya, dad, best, story, **kids**, video, collection, whole, **family**, popular, **friend**, **play**, **toy**, etc.

→ Is this video made for kids?

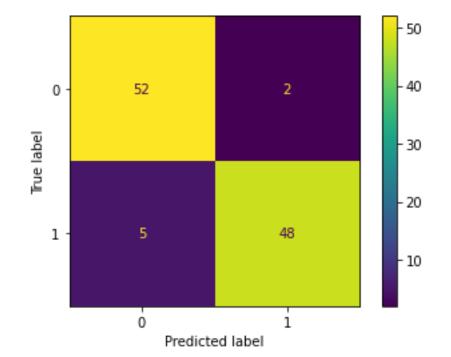
Naïve Bayes Classifier – Training

- Only 5% of all videos are made for kids
 - In order to handle class imbalance, the same number of non-kids videos (5%, or around 200 videos) is used for training

- Data is split further
 - 75% of videos are assigned for training
 - 25% of videos are reserved for testing

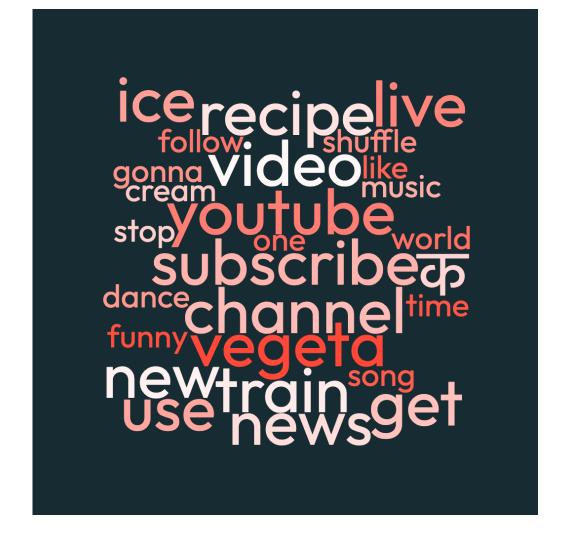
Naïve Bayes Classifier – Result

- 91% of videos "made for kids" are identified accurately
- 96% of videos "not made for kids" are not flagged



Naïve Bayes Classifier – Top Words for Kids Videos

Naïve Bayes Classifier – Top Words for Non-Kids Videos



II. CONCLUSION

Recommendation & Key Takeaways

- Use a Naïve Bayes model
 - 1. It has high prediction accuracy
 - Words such as kid, song, and miraculous are highly associated with kids videos

2. The model can save manual human efforts in identifying videos "made for kids"

Next Steps

• Include more videos for training and testing

Potentially use comments as an additional data point

Thank You

Sue Lim

- Email: <u>sue.suyeonlim@gmail.com</u>
- GitHub: https://github.com/suesuyeonlim
- LinkedIn: https://www.linkedin.com/in/suelim/