

# Identifying YouTube Videos “Made for Kids”

Sue Lim  
July 21, 2023

# I. PROJECT BACKGROUND

# YouTube Videos “Made for Kids”

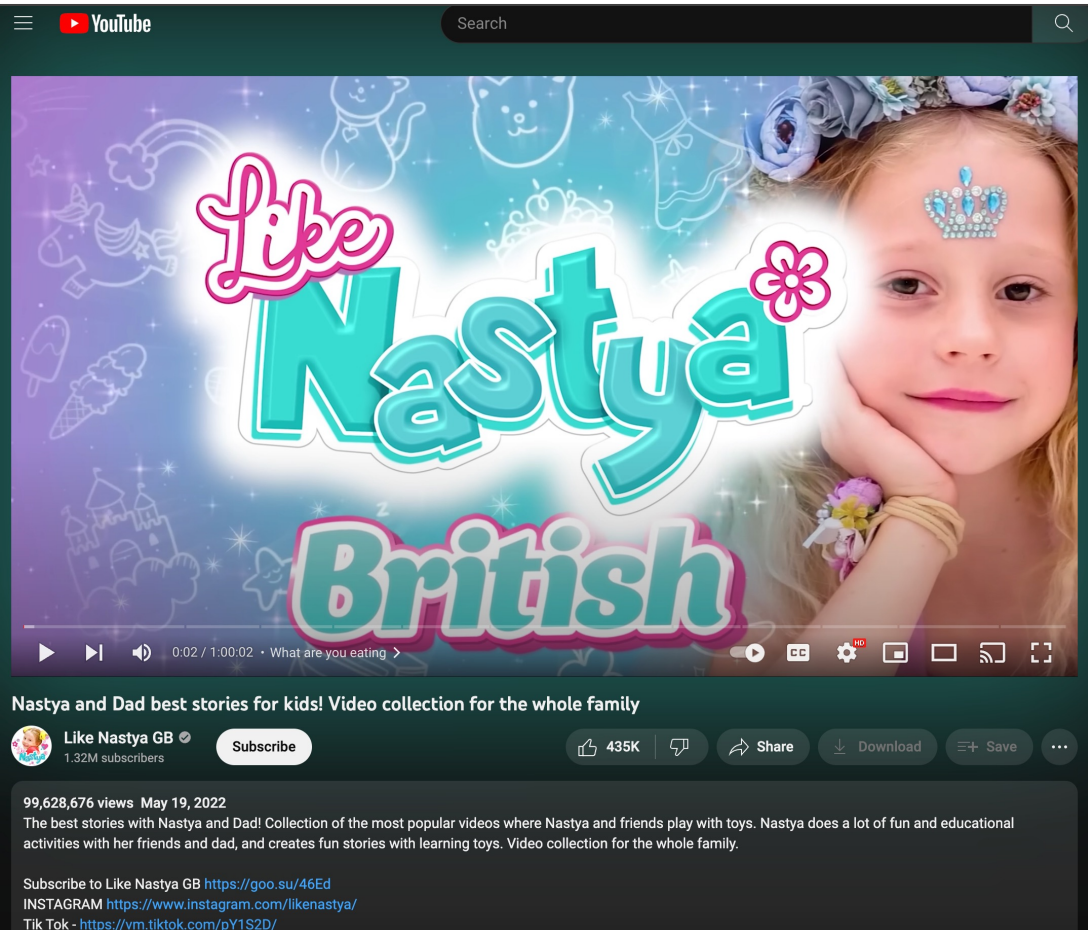
- FTC requires identifying YouTube videos “made for kids” under COPPA (Children’s Online Privacy Protection Act)
  - Such videos are under certain restrictions
  - YouTube may face **legal consequences** if content is not categorized correctly
- ➡ **The need to flag videos correctly**

The screenshot shows the YouTube video settings interface, specifically the 'Details' tab. At the top, there are three tabs: '1 Details' (active), '2 Video elements', and '3 Visibility'. Below the tabs, the 'Audience' section is visible. It asks 'Is this video made for kids? (required)' and provides a detailed explanation: 'Regardless of your location, you're legally required to comply with the Children's Online Privacy Protection Act (COPPA) and/or other laws. You're required to tell us whether your videos are made for kids. [What's content made for kids?](#)'. There are two radio button options: 'Yes, it's made for kids' (which is selected) and 'No, it's not made for kids'. A grey information box is present next to the 'Yes' option, stating: 'Features like personalized ads and notifications won't be available on videos made for kids. Videos that are set as made for kids are more likely to be recommended alongside other kids' videos. [Learn more](#)'. At the bottom, there is a collapsed section for 'Age restriction (advanced)'.

# 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-mRuy1JLQQqEk',
'items': [{ 'kind': 'youtube#video',
'etag': '1EXzyGvBtQsYeIhPTWILufbNAXE',
'id': 'UQyHkp8-B6E',
'snippet': { 'publishedAt': '2022-05-19T09:00:11Z',
'channelId': 'UC24uDUHbhwJTM_QPaUke5mI',
'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 friends play with toys. Nastya does a lot of fun and educational activities with her friends and dad, and creates fun stories with learning toys. Video collection for the whole family.\n\nSubscribe to Like Nastya GB https://goo.su/46Ed\nINSTAGRAM https://www.instagram.com/likenastya/\u200b\nTik Tok - https://vm.tiktok.com/pY1S2D/',
'thumbnails': { 'default': { 'url': 'https://i.ytimg.com/vi/UQyHkp8-B6E/default.jpg' },
'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'],
'categoryId': '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 friends play with toys. Nastya does a lot of fun and educational activities with her friends and dad, and creates fun stories with learning toys. Video collection for the whole family.\n\nSubscribe to Like Nastya GB https://goo.su/46Ed\nINSTAGRAM 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': { 'totalResults': 1, 'resultsPerPage': 1}}
```

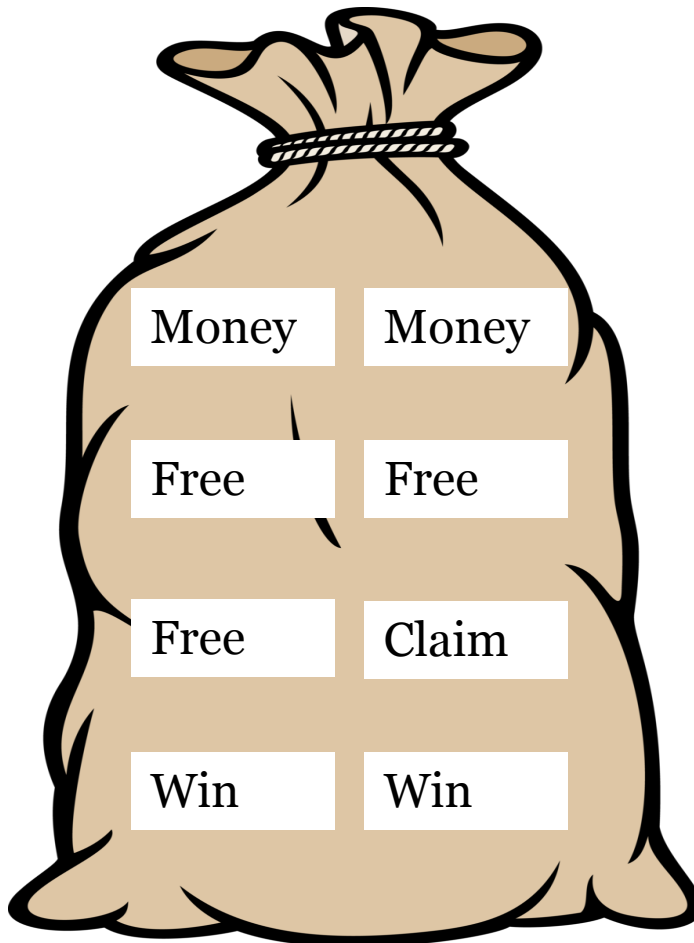
# Benefits of Machine Learning

- Highly scalable
- Reduce manual human efforts
- Consistent criteria in identifying videos

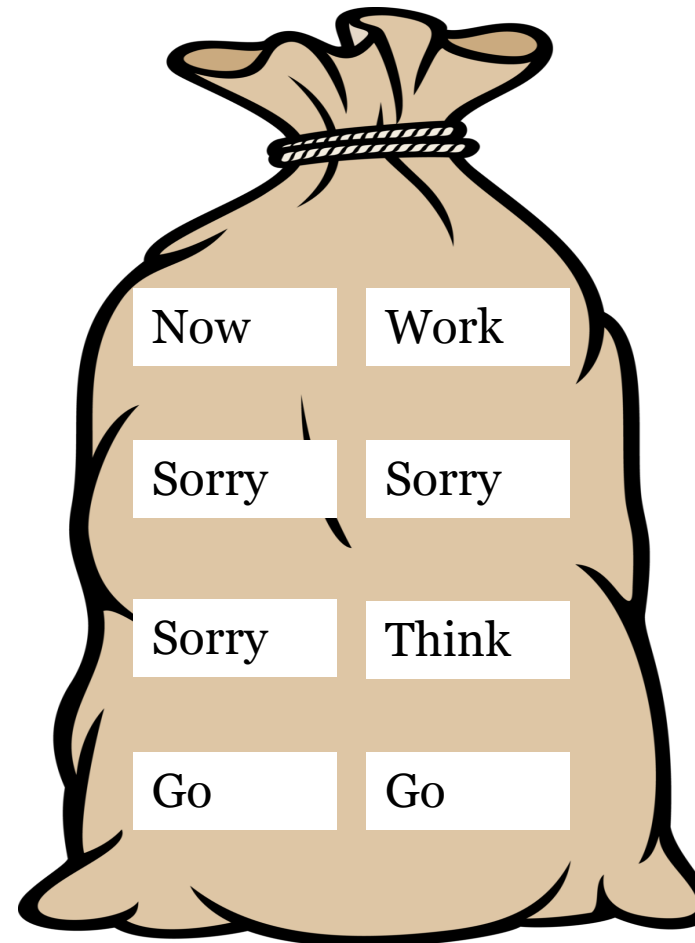
## II. MAIN MODEL

# Naïve Bayes Classifier

Spam



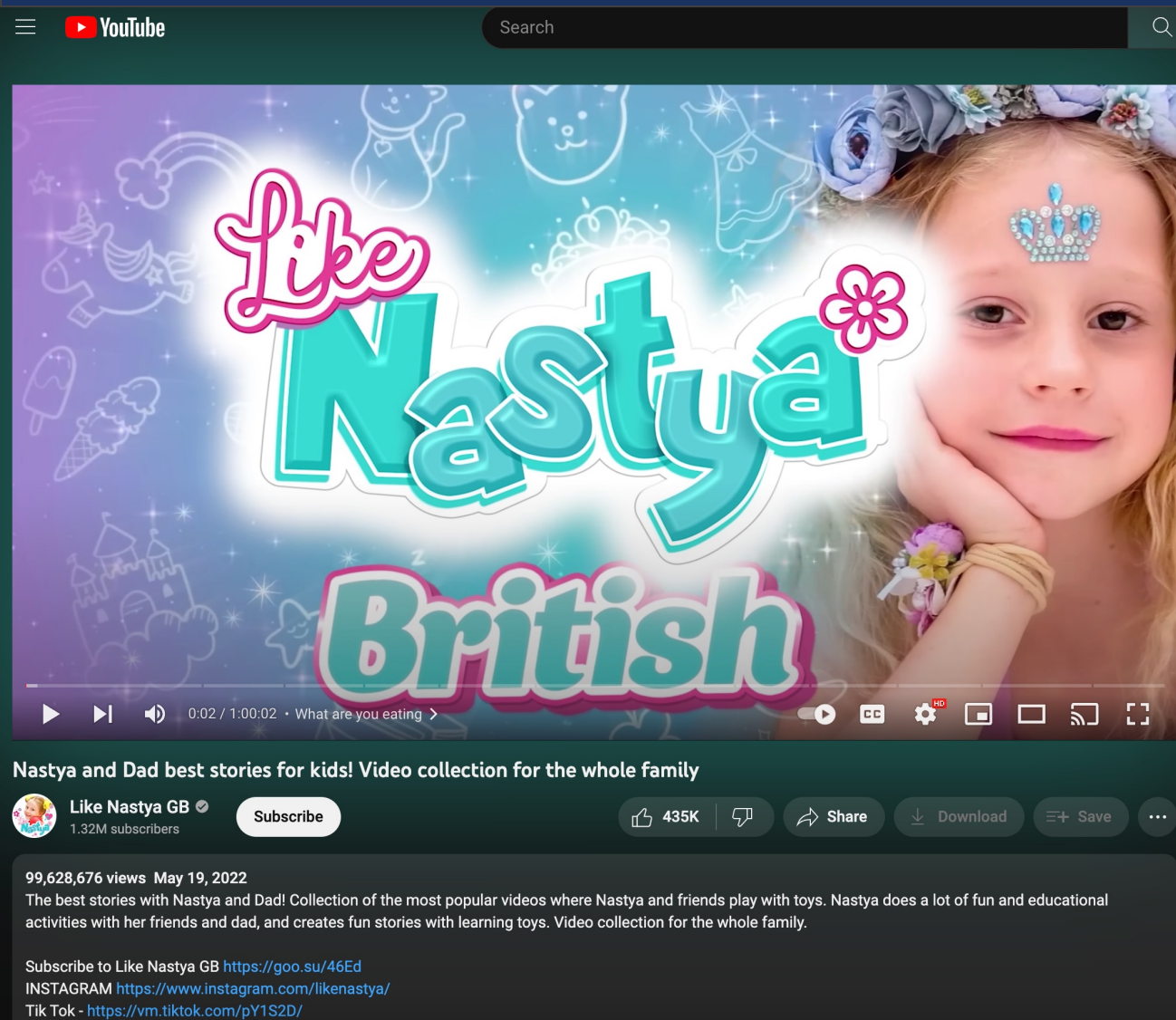
Ham



**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
- **94%** of videos “not made for kids” are not flagged

	Kids	Non-Kids
Predicted as Kids	48 ( <b>91%</b> )	51 ( <b>94%</b> )
Predicted as Non-Kids	5 (9%)	3 (6%)
Total	53 (100%)	54 (100%)

# 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: [sue.suyeonlim@gmail.com](mailto:sue.suyeonlim@gmail.com)
- GitHub: <https://github.com/suesuyeonlim>
- LinkedIn: <https://www.linkedin.com/in/suelim/>