

Progress presentation

Enhancing the NPO Recommendation System with Metadata
Eileen Kapel

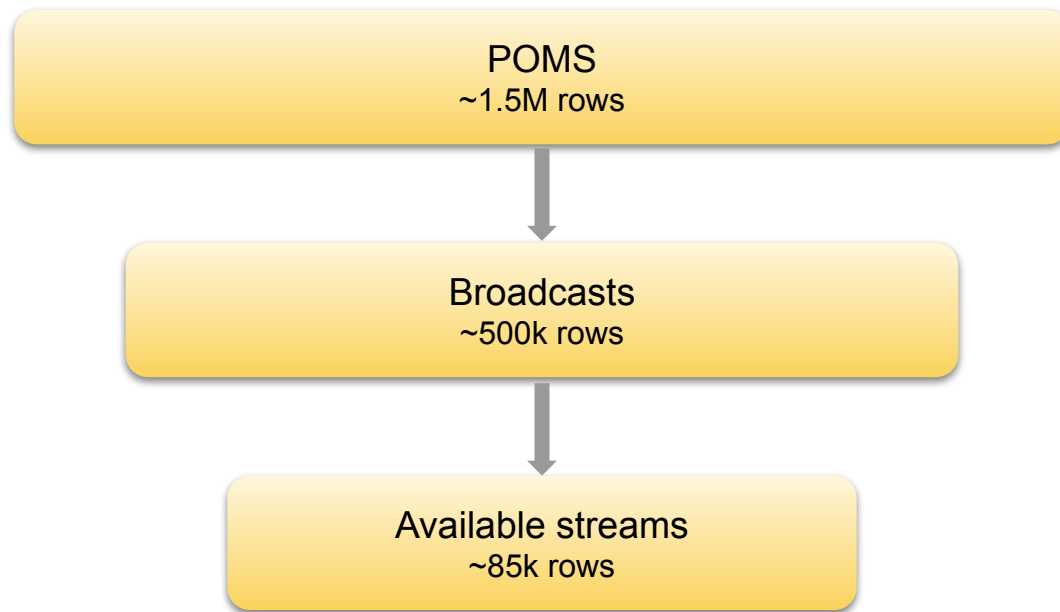
1. Research questions

*‘Can **content-based filtering** using **metadata** enhance the current **recommendation system** of NPO Start which uses **collaborative filtering**?’*

1. Research questions

- Sub-questions
 - What is the current collaborative filtering performance?
 - How much can the recommendation system be improved by using different content-based filtering methods integrated with the current system?
 - Which metadata features improve the recommendation the most?

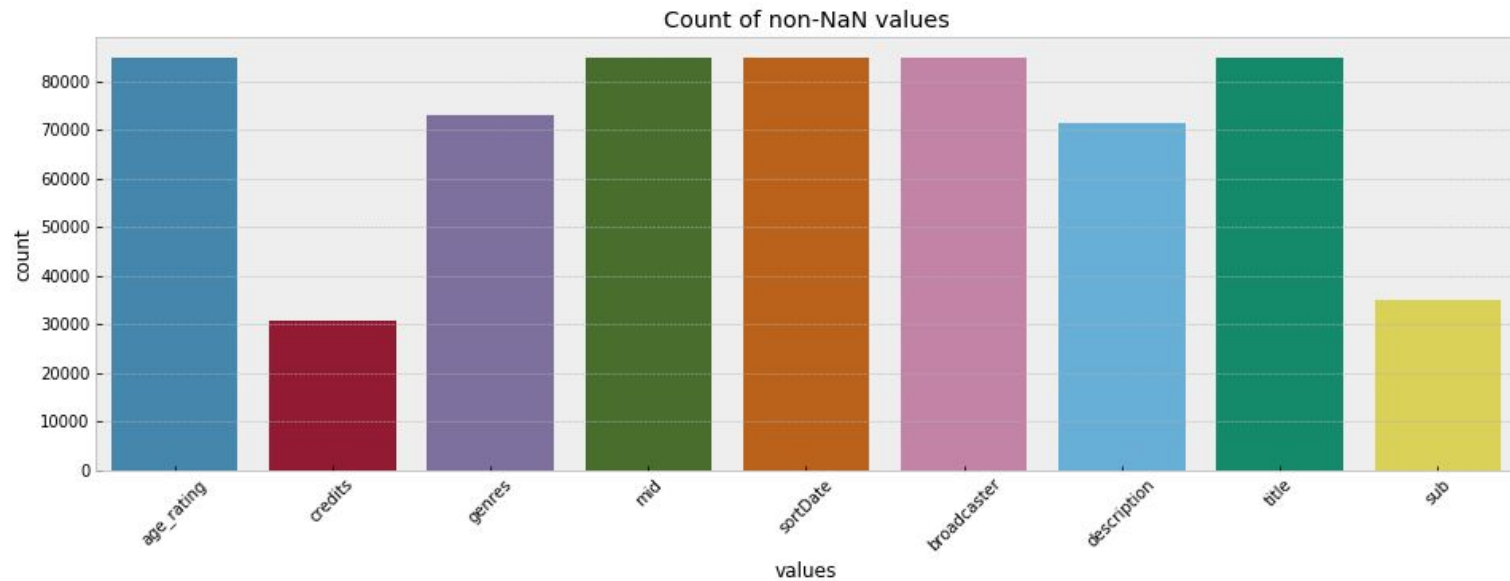
2. The data



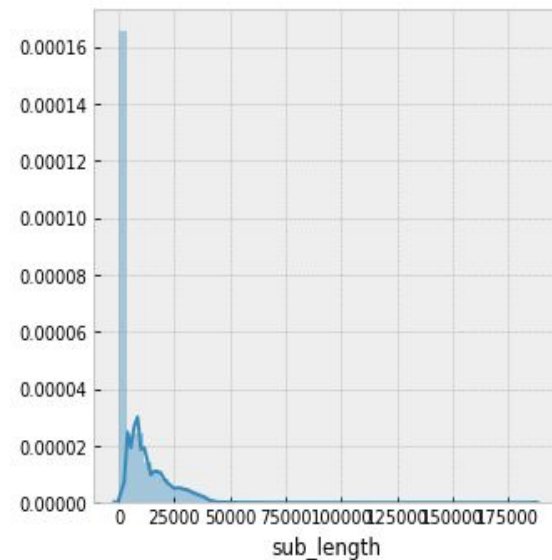
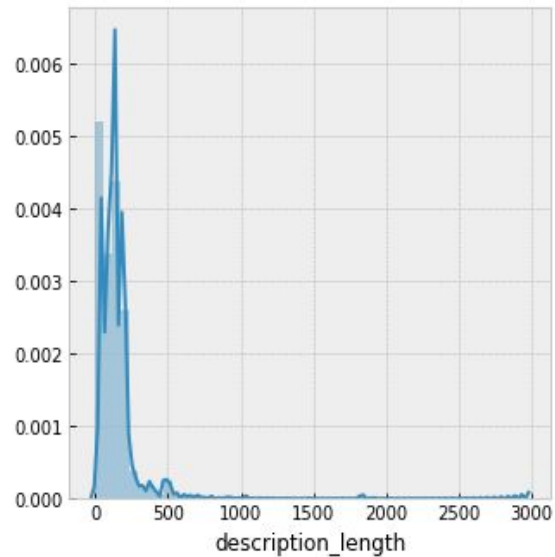
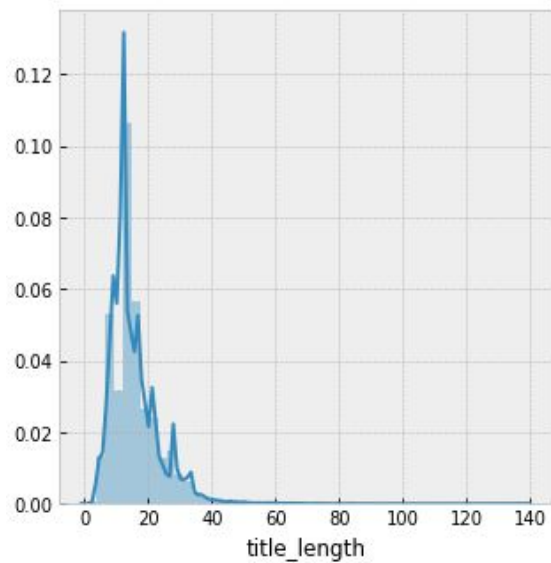
2. The data

| age_rating | credits | genres | mid | sortDate | broadcaster | description | title | sub |
|------------|--|--|--------------|-------------------------------|-------------|---|-------------------|---|
| ALL | [(Freriks, Philip, PRESENTER), (Rossem, Maarte...] | [(3.0.1.6, [Amusement])] | NCRV_1616517 | 2013-08-06 18:25:00 UTC | NCRV | Kennisquiz met een knipoog. | De slimste mens | 888 wie kent de steden van europa... en weet w... |
| ALL | None | [(3.0.1.1, [Jeugd]), (3.0.1.7.21, [Informatief...] | POW_00421885 | 2012-03-31 16:45:00 UTC | NOS | None | NOS Jeugdjournaal | None |
| ALL | [(van de Westelaken, Rik, PRESENTER)] | [(3.0.1.7.21, [Informatief, Nieuws/actualiteit...] | AT_2096527 | 2018-07-02 16:25:00 UTC | AVTR | O. a.: Milieudefensie roept banken op te stopp... | EenVandaag | 888 dit is eenvandaag. goedenavond. hij is al ... |

2. The data



2. The data



3. Achievements

| Week | Achievements |
|------|--|
| 1 | Introduction section done. Pre-processed data. |
| 2 | EDA and data section done. Investigated current performance. |
| 3 | Literature read. Literature section done. |

4. Deliverables

| Week | Achievements |
|------|--|
| 4 | Baseline algorithm was set up and evaluated. |
| 5 | Investigated integrating content-based filtering and collaborative filtering. Method section done. |
| 6 | <i>Mid-presentation.</i> Hybrid recommender system was set up. |
| 7 | A/B testing was set up. |
| 8 | Checked results of A/B testing. Experimented and tuned hyperparameters. |
| 9 | Checked results of A/B testing. Experimented and tuned hyperparameters. |
| 10 | Results section done. |
| 11 | Conclusion & discussion section done. |
| 12 | Thesis done. |

5. The internship

- Data
 - Access
 - Correct subset
- Supervisors