# Happy News Retrieval

~ Source of Your Joy


- Aayush Gautam
- Anuska Pant
- Rohit Sah
- Yatin Gupta


1. **What is exactly the function of your tool? That is, what will it do?**
   Our tool will retrieve “happy” news in real time from news publishing sites like google news and social media networks like Reddit, Twitter etc. The user will be provided with happy news based on their profile. They can also query for specific topics that they want to retrieve happy news from. The recommendation system will keep on improving and personalize its recommendations based on the implicit feedback received from the user through queries and clicks.

2. **Why would we need such a tool and who would you expect to use it and benefit from it?**
   In the modern world where social media has become the main source of information for many people, it is important to assure that it does not become a place of just communicating anger and hatred. A lot of things happen in the world, and research has shown that negative and inflammatory information tends to spread the fastest across the web. With our tool, we hope to bring a balance by showing happy news to people who surf the internet.

   It could be useful for individuals who are looking for positive news or uplifting content on the internet. Our tool will provide a quick and easy way for a user to feel good.
   Some potential users of this system could include: People who want to balance out negative news, Individuals who are struggling with mental health, News junkies who want to stay informed.

3. **Does this kind of tool already exist? If similar tools exist, how is your tool different from them? Would people care about the difference? How hard is it to build such a tool? What is the challenge?**

   There are various people and groups around the world who are working on making the web a happier place. Some of them are as follows:
   a. Good News Network: The website is an archive of 21,000 positive news stories from around the globe. GNN was founded in 1997 because the media was failing to report positive news. In the 1990s while homicide rates in the U.S. plummeted

by 42 percent, television news coverage of murders surged more than 700%, according to the Center for Media and Public Affairs.

b. SomeGoodNews: Youtube channel by John Krasinski : "a news show dedicated entirely to good news":
c. Upworthy: A social impact company with a mission to empower people to be a force of good.

All of these are platforms that research and publish novel good news. Our plan is to build a tool that classifies real time information into good news categories, and retrieve relevant articles based on the user's query in a quick and easy way.

There are various tools that perform analysis of different affective states expressed in the text on the web. And the web has many search engines: Google, Bing etc.

We plan to bring together the power of deep learning to analyze text and the technique of information retrieval to deliver happy news to people.

4. **How do you plan to build it? You should mention the data you will use and the core algorithm that you will implement.**

For data, we will use Twitter and Reddit. Twitter has already defined happy news channels like @goodable. Reddit also has similar subreddits like r/wholesomememes. These channels can act as positive data for us so that we need not manually annotate. We will take this data and combine with other channels to create a training dataset.

Next, we will train and use the Bert model for sentiment analysis. This will form a basis for us to solve the “cold start” problem and start classifying live tweets and threads to recommend to users.

Then we will continue to improve our recommendation engine by recording and using the implicit feedback (clicks) and creating a global profile of our users. As the data increases, we will slowly transform it from a global perspective to a user profile for personalized recommendations.

The main tool will be deployed as a slack bot and users would be able to interact with slack slash commands like “/happynews”.

5. **What existing resources can you use?**

a. For filtering good news from regular news content, we will use the sentiment analysis approach. We will finetune a BERT-based model using happy news curated from Twitter channels like @goodable.

b. We will use the embeddings from the BERT model trained in the English language for content representation in news articles for content-based recommendation.
c. For recommendations using implicit and explicit feedback, we will use matrix factorization and BPR-based techniques taught in the class. Finally, we will combine the content-based recommendation score and collaborative filtering score using weighted average techniques.
d. For data, we can use the news from “Good News Network” to evaluate our sentiment analysis tool. Available APIs for Twitter, Reddit, google news, etc. will be used to download information for datasets and real-time evaluation.

6. **How will you demonstrate the usefulness of your tool?**

   To evaluate the performance of the model we will download actual good news from sources like goodnewsnetwork, google news etc and classify news into happy or not, to get metrics like accuracy, precision etc.
   We will deploy the slackbot to our class slack workspace before our project presentation so we can gather data on its usefulness from our friends. They can help us gather data on whether the news they saw made them feel happy or not. We also plan to deploy it as a slack bot which teams and companies can use to promote positivity in the workplace.

7. **A rough timeline to show when you expect to finish what. List a couple of milestones.**

| Week | Description |
|---|---|
| 1 | Configure data sources: Creating training and test datasets. Build API wrapper around Twitter and Reddit to extract live news from them |
| 2 | Build Sentiment Analysis Tool to classify news into “happy” category |
| 3 - 4 | Build a realtime recommender engine - classify real time news into happy categories and tag them with specific query words like “layoff”. Deploy as a slack bot. |
| 5 | Build feedback loop from implicit actions to improve recommendations |
| 6 | Switch from global to personalized recommendations |
| | |