Personalized News Recommendation

Sai Ashritha Kandiraju
skandira@ucsd.edu

Prajwala Thatha Manjunatha
pthatham@ucsd.edu

Abstract

Online news websites such as Microsoft News, Google news, BBC, CNN, NYTimes are a popular source of news for a large number of users. Data explosion on the internet results in several news articles being generated and uploaded to the web every single day. Due to the plethora of information, it becomes increasingly difficult for a user to find news relevant to the user's taste. Personalized news recommendation can help a user navigate the information overload by recommending relevant news articles and improving the user's experience. Compared to product and movie recommendations which have been exhaustively surveyed, news recommendation has only gained traction over the last few years. The limited research in news recommendation is mainly due to the lack of a high-quality dataset which has now been alleviated by MIND - Microsoft News Dataset [9]. In this report, we present recommendation techniques for personalized news recommendation on Microsoft News Dataset (MIND).

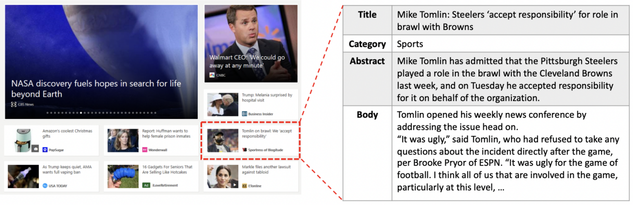
Keywords: News Recommendation, Collaborative-Filtering, Content-Based Filtering, FastFM, Popularity

1 Data Analysis

This section describes the dataset used for the assignment and the exploratory analysis conducted on the data. This section also highlights some interesting findings from the data.

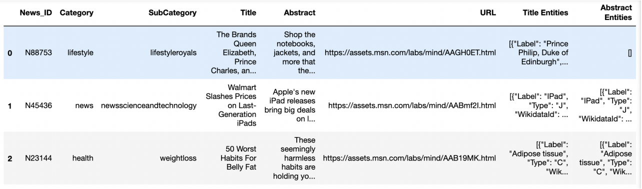
1.1 Dataset

The Dataset used for our recommendation task is Microsoft News Dataset (MIND). This dataset was constructed by Microsoft by collecting user behavior logs of Microsoft News. For building the dataset, they have sampled 1 million users, anonymized them and recorded their click behavior during a time period of 6 weeks from October 12 to November 22, 2019. The train set is data pertaining to the 5th week based on the modeled interactions of the user over the past 4 weeks. The dataset contains impression log records for each user that contains information about the clicked news articles, non-clicked news articles and historical click data of this user as shown in Fig. 3. Each row in the file *behaviours.tsv* from the dataset contains a timestamp of when the impression log has been recorded. Each news article in the data is identified by rich contextual features like news title, abstract, body, a category label like "Sports", a sub-category label like "Football_NFL" as represented in Fig. 1 and Fig. 2. Each news article is also associated with entities from Wikidata. The embeddings from these entities and their relations are also included in the MIND dataset. Thus, the user news interaction impression logs serve as a rich source of implicit feedback from the user.


NASA discovery fuels hopes in search for life beyond Earth
CBS News
Walmart CEO: 'We could go away at any minute'
CNBC
Trump: Melania surprised by hospital visit
Business Insider
Amazon's coolest Christmas gifts
PopSugar
Report: Huffman wants to help female prison inmates
Wonderwall
Tomlin on brawl: We 'accept responsibility'
Sportress of Blogitude
As Trump keeps quiet, AMA wants full vaping ban
USA TODAY
16 Gadgets For Seniors That Are Selling Like Hotcakes
Ad iLoveRetirement
Markle files another lawsuit against tabloid
ETonline

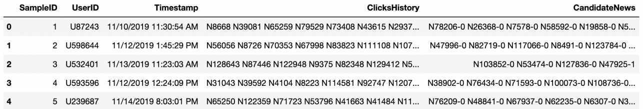
Title	Mike Tomlin: Steelers 'accept responsibility' for role in brawl with Browns
Category	Sports
Abstract	Mike Tomlin has admitted that the Pittsburgh Steelers played a role in the brawl with the Cleveland Browns last week, and on Tuesday he accepted responsibility for it on behalf of the organization.
Body	Tomlin opened his weekly news conference by addressing the issue head on. "It was ugly," said Tomlin, who had refused to take any questions about the incident directly after the game, per Brooke Pryor of ESPN. "It was ugly for the game of football. I think all of us that are involved in the game, particularly at this level, ...

Figure 1. An example of a news article from Microsoft News Homepage and its representation in the dataset



	News_ID	Category	SubCategory	Title	Abstract	URL	Title Entities	Abstract Entities
0	N88753	lifestyle	lifestyleroyals	The Brands Queen Elizabeth, Prince Charles, an...	Shop the notebooks, jackets, and more that the...	https://assets.msn.com/labs/mind/AAGH0ET.html	[{"Label": "Prince Philip, Duke of Edinburgh",...	[]
1	N45436	news	newsscienceandtechnology	Walmart Slashes Prices on Last-Generation iPads	Apple's new iPad releases bring big deals on l...	https://assets.msn.com/labs/mind/AABmf2I.html	[{"Label": "IPad", "Type": "J", "WikidataId": ...	[{"Label": "IPad", "Type": "J", "WikidataId": ...
2	N23144	health	weightloss	50 Worst Habits For Belly Fat	These seemingly harmless habits are holding yo...	https://assets.msn.com/labs/mind/AAB19MK.html	[{"Label": "Adipose tissue", "Type": "C", "Wik...	[{"Label": "Adipose tissue", "Type": "C", "Wik...

Figure 2. News article representation in MIND dataset



	SampleID	UserID	Timestamp	ClicksHistory	CandidateNews
0	1	U87243	11/10/2019 11:30:54 AM	N8668 N39081 N65259 N79529 N73408 N43615 N2937...	N78206-0 N26368-0 N7578-0 N58592-0 N19858-0 N5...
1	2	U598644	11/12/2019 1:45:29 PM	N56056 N8726 N70353 N67998 N83823 N111108 N107...	N47996-0 N82719-0 N117066-0 N8491-0 N123784-0 ...
2	3	U532401	11/13/2019 11:23:03 AM	N128643 N87446 N122948 N9375 N82348 N129412 N5...	N103852-0 N53474-0 N127836-0 N47925-1
3	4	U593596	11/12/2019 12:24:09 PM	N31043 N39592 N4104 N8223 N114581 N92747 N1207...	N38902-0 N76434-0 N71593-0 N100073-0 N108736-0...
4	5	U239687	11/14/2019 8:03:01 PM	N65250 N122359 N71723 N53796 N41663 N41484 N11...	N76209-0 N48841-0 N67937-0 N62235-0 N6307-0 N3...

Figure 3. An example of a user news interaction impression log in MIND dataset

1.2 Exploratory Data Analysis

The dataset contains 2,186,683 samples in the training set, 365,200 samples in the validation set, and 2,341,619 samples in the test set. The statistics of the dataset are summarized in Table 1.

From exploratory analysis we figure out that the interaction data is heavily skewed with respect to non-clicked data. About 96% of the candidate news (recommended articles) are not clicked by user. This distribution is plotted in Fig. 4. This suggests that accuracy might not be the best evaluation metric for our predictive task as a trivial predictor that always predicts 0 reports an accuracy of 96%. In the user clicked news articles also we observe few dominant categories like "Sports" and "News" as shown in Fig. 5. We then explore the distribution of categories and subcategories

Table 1. Statistics of Train set and Entire Data Set

Statistics	Train Set	Entire Dataset
#News	101527	161,013
#Users	711222	1,000,000
#News Categories	18	20
#News Sub Categories	285	285
#Impressions News	10,507,374	15,777,377
#Clicks News	16,675,933	24,155,470
#Entities News	1,893,221	3,299,687
Avg title length	10.31	11.52

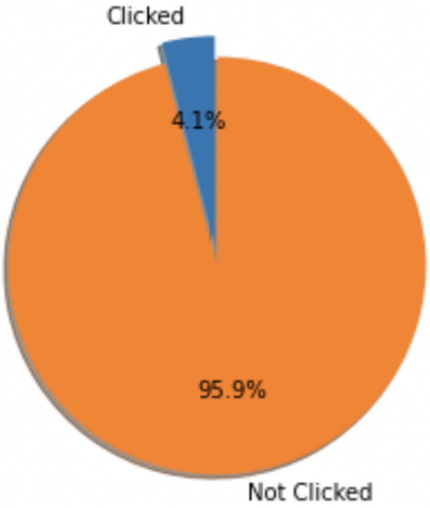


Figure 4. Distribution of Clicked and Not Clicked data

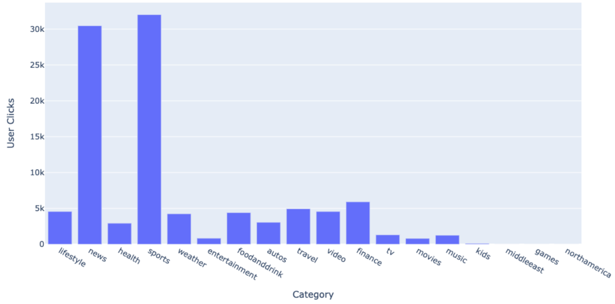


Figure 5. Distribution of User Clicked News Categories

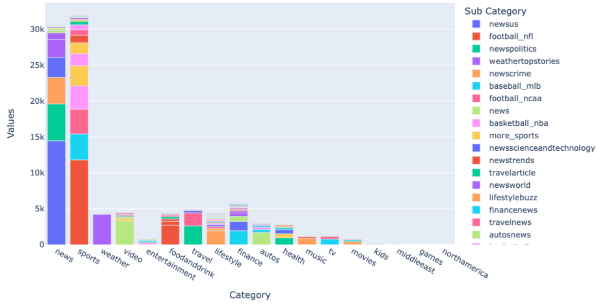


Figure 6. Distribution of Categories and Sub-Categories among News Articles

among the new articles. This histogram is plotted in Fig. 6. From these distributions it is evident that the data is rich in linguistic characteristics that can help us determine news representations for news modeling. These plots also convey that popularity could be a good feature for this dataset as the user clicks are skewed towards popular categories like "Sports" and "News". This makes sense as people tend to click on trending or popular news while browsing through news articles. An interesting point to note here is that during the week when the train set was modeled (10th - 14th Nov 2019), it was the NFL season which explains the popularity with respect to sports news.

Fig. 7. shows the plot of user clicks over time. Since the user interaction data has been modeled over a week, the plot shows the distribution of clicks over 6 days. MIND dataset samples the 7th day of the week as the validation data. Fig. 8. shows the survival time distribution among news articles presented in the MIND dataset paper [9]. In the paper, survival time is estimated as the time interval between the first and last appearance of a news article in the dataset. The plot from the paper illustrates that the survival time of more than 84.5% news articles is less than 48 hours. This can be explained by the nature of news over the web. Trending and exciting news articles quickly replace other news articles and this cycle repeats. Thus it can be concluded that news articles are short-lived and get out-of-date quickly.

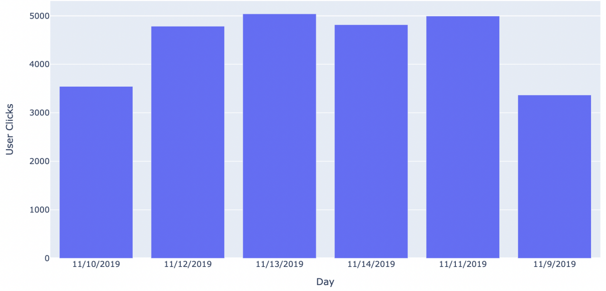


Figure 7. Distribution of Clicks over a week

2 Predictive Task

This section outlines the objective of the predictive task, the methodologies to evaluate the model at this predictive task, baselines used for validating the predictions, description

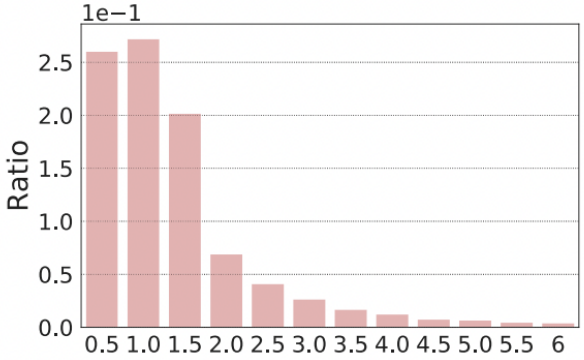


Figure 8. Survival time of news articles in MIND Dataset

of data pre-processing, feature engineering and extracted, how the exploratory data analysis in Section 1 justifies the features extracted and evaluation metrics.

2.1 Task Description

The objective is to build a news recommendation system which predicts news articles that will be clicked by the user from a given set of candidate news articles. We aim to identify news articles that a user shows interest to read about from a relatively large set of candidate news articles. Given a user *u*, timestamp of recommendation *t*, historical news click behaviors of user *u* as *H*, candidates news articles *C*, the task is to predict whether each of the candidate news articles will be clicked or not *i.e*: for each article *c* in *C*, a binary label *0,1* is based on whether or not the user *u* clicks on *c*.

2.2 Baselines

Here are the baselines listed to compare and validate the performance of our approaches.

- Trivial predictor - predicts not-clicked for every candidate news article.
- Most popular category - predicts clicked for news articles in most popular categories(40th) percentile.
- Most popular per day category- predicts clicked for news articles in most popular categories(40th percentile) in that day.
- Ablation methods:
  - Most recently clicked - predicts clicked only for most recently clicked news articles for each user
  - Most recently clicked - predicts clicked for news articles that belong to the category of the most recently clicked news articles for each user.
  - News text similarity - predicts clicked if news title/text similarity between candidate news articles and articles in click history of user exceeds a particular threshold.

2.3 Data preprocessing

The dataset included training samples where the historical news clicks behavior of certain users was not available (marked *nan*). All such samples were discarded while constructing training samples using just the click history but were considered while building training samples using the impression logs.

2.4 Evaluation metrics

Various metrics were analyzed for evaluating our model and the final model alternatives were evaluated based on AUC and BER. The choice of these metrics is justified as follows:

- Accuracy - Accuracy proved to be a non-informative metric for this use-case since 96% of the dataset reveals non-clicks (*0 label*). Thus, trying to optimize our model based on accuracy will drive the model towards trivial prediction *ie:* predicting *0* everywhere.

- Balanced Error Rate (BER) - Balanced Error Rate proves to be a useful metric since it cannot be optimized by trivial predictions like in the case of accuracy. A trivial prediction will result in a BER of *0.5*.

- Area under Curve (AUC) - Area under the ROC curve also proves to be a useful metric since it optimizes the predictions based on both True Positives and False Positives, which will simultaneously increase if we decrease the threshold for clicks detection.

3 Method

This section describes the final model adopted for news recommendation on the MIND dataset in detail along with various models and news recommendation approaches that were attempted, reasons for choosing these models, justification of feature representations, and strengths and weaknesses of the above models.

3.1 Attempted news recommendation methods

This subsection discusses the various news recommendation methods that were implemented on the MIND dataset in different settings. The choice of model alongwith strengths, weaknesses and results are described in detail.

3.1.1 Similarity based recommendation The first recommendation method attempted is the Jaccard similarity heuristic to recommend candidate news articles. This is the simplest approach to the recommendation problem, where a candidate news article is predicted to be clicked by the user if it is similar to news articles that have been previously clicked by the user. This is a memory based approach and an AUC of 0.52 with BER of 0.49 was obtained on the MIND validation set.

There are multiple problems to this approach for the news recommendation problem:

- Sparsity of the user and article interaction sets - Each user only would have clicked a small subset of candidate news articles in the past and hence the user and item sets for Jaccard similarity will be sparse and hence would not work very well.
- Cold start problem - For new users or new news articles, there is no interaction history present based on which predictions can be made. To handle cold start for new users, we recommend news articles from most popular categories. For new items, we find the text similarity between news articles previously clicked by the user and the current news article and threshold it to assign label 0 or 1.
- The Jaccard similarity heuristic does not scale well to increased number of interactions being added in real-time in a practical scenario since it needs to be constantly retrained being a memory based approach.

3.1.2 Popular category per user based filtering This recommendation method predicts a news article is clicked by a user if it belongs to any of the most popular news categories clicked previously by that user. Here, the entire historical news clicks behavior (over 5 weeks) is considered while constructing the most popular news categories for each user. The threshold for determining the popular categories is a hyper-parameter tuned on the MIND validation dataset. We obtained 40th percentile to be the best threshold. The AUC obtained was 0.573 with a BER of 0.426 on the MIND test set.

This recommendation method beats the trivial predictors as it considers the historical clicks behavior of the user to obtain the most popular categories. Thus, it still qualifies as a personalized news recommendation system and works well to a good extent due to the skewed nature of the dataset. It is however not the best model since it does not consider the timestamps of the clicks. The model is not sensitive to changing preferences of the user over time *i.e:* it weights his interests built over a long period of time greater than the most recent interests of the user, which is not the best recommendation method always considering the low survival time distribution of news articles.

3.1.3 Bayesian Personalized Ranking (BPR) This recommendation method adopts a ranking scheme by assigning scores to news articles such that clicked news articles are ranked higher relative to the non-clicked news articles. This method accounts for implicit feedback where news articles that are not clicked are not necessarily assumed as negative interactions as in case of logistic regression or classification, since the items that are not clicked are the candidate news articles that need to be recommended to the user. In other words, BPR is a pairwise predictor model where neither news article is given a negative label, but the clicked news articles are given more positive scores (thus, lower rank) in comparison to the non-clicked ones.

Table 2. Features for Logistic Regression

Feature Name	Encoding	Dimensionality
date	one-hot	6
news_title	one-hot	1
news_category	one-hot	18
news_sub_category	one-hot	84
news_title_100	one-hot	100
user_history	float	tf-idf
entities	one-hot	100

We implemented BPR using the *Implicit* python library. For this, we build a sparse interaction matrix and store only the articles that users interacted with as clicks as positive examples. We then use *implicit* to build the BPR model with MIND train set. For each user-item interaction validation sample, we obtain the most recommended items for each user and most similar items to each item. If either the most recommended items contains the item or the any of the most similar items obtained have been previously clicked by the user, we assign a label of 1 else 0. The number of latent factors was tuned over the MIND validation test and best results were obtained with factors = 5. The AUC obtained was 0.54 with a BER of 0.48 for the MIND test set.

This method did not work as the best model for the task as implicit data is inherently noisy. The candidate news articles that are not clicked are inherently ranked lower and hence there is still some amount of negativity associated with the implicit signals.

3.1.4 Logistic Regression This recommendation method employs the use of manually extracted features to classify a news article as "Clicked" and "Not Clicked". We train a logistic regression model with the selected features to predict the recommended news article. We experimented with several features to model users and news articles like user history, timestamp, news title, news abstract, news category, news subcategory, popularity of a news article, user-user similarity, item-item similarity etc. We tried to learn text representations from news articles using techniques like BoW, TF-IDF, LDA. After conducting a series of experiments we have finalized our features and are described in the Table 2 .

Most of the above features have been one-hot encoded due to their categorical nature. For the news article titles, we tokenized the titles, removed stop words from the tokens, Lemmatized the words using WordNet lemmatizer from spacy, Detected entities using spacy and combined them with entities embedded in MIND news data. We then picked the top 100 entities based on occurence across news articles, similarly we also picked the top 100 unigram words and generated one-hot encoded features. The feature importance for the top 100 unigram words is shown in Fig. 9.

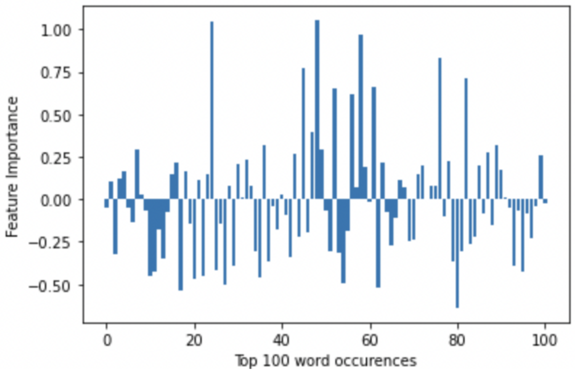


Figure 9. Feature importance

3.2 FastFM Model

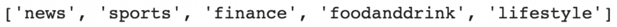
The models that we considered so far are either completely built on interaction data (Similarity, Popular category per user, BPR) or only using features extracted from metadata or problem structure (Logistic Regression). However, in practice, we require more complex models that are built on interaction data and also leverage side-information and features extracted from the problem structure to learn more complex patterns. Motivated by the above, we employed the factorization machine approach to solve the news recommendation problem. The factorization machine extends the latent factor models by incorporating features in addition to pairwise interaction data.

We implement the factorization machine recommendation approach using *fastFM*[2] library. The interaction matrix includes user encoding, item encoding along with one-hot encoded features - news categories, news sub-categories, length of news titles one-hot encoded by converting the lengths to categorical range variables. Text representations using bag of words model for the news titles, abstract could be leveraged but one-hot encoding such features would result in very long feature vectors and hence we limited our features to news categories and sub-categories.

The model was trained on MIND train dataset with rank=5 used for the second order interactions, 1000 iterations, and regularization parameters of 0.1 and 0.5 for the coefficients of linear combination and coefficients for pairwise coefficients. The hyper-parameters were tuned on the MIND validation dataset.

We next address the cold-start problem for both new users and new news articles which can be quite common in the test set. Prior to handling the cold start problem, we skipped all the new users and new news articles that were not seen earlier in the MIND train set. This resulted in an AUC of 0.97 with BER of 0.04 which clearly showcased a case of recommendation for only seen users and articles in the train set.

To address the cold start problem for new users, we recommend articles from most popular categories (60$^{\text{th}}$ percentile) computed from the MIND train set. To tackle cold start for new news articles, we compute similarity between the categories of articles previously clicked by the user and the current article and predict 0 or 1 according to a threshold (0.3). The FastFM [2] model returned AUC of 0.662 and BER of 0.338 on the MIND test set, proving to be the best of all the news recommendation approaches attempted at the MIND dataset. Fig. 10 shows the most popular categories obtained for handling the cold start problem.



```
['news', 'sports', 'finance', 'foodanddrink', 'lifestyle']
```

Figure 10. Most popular categories used for Cold Start in FastFM

4 Literature and Research

4.1 Dataset Literature

As mentioned in the data analysis section, MIND dataset used in our recommendation task is relatively recent. Motivated by the lack of a publicly available high-quality benchmark dataset for news recommendation, Microsoft research group has contributed MIND dataset to the community. Due to the public availability of a large dataset with millions of interactions and rich textual features, mostly deep learning based recommender systems have been a popular choice for this dataset. A few popular datasets that were employed for news recommendation prior to the availability of MIND dataset are Plista[5], Adressa[4], Yahoo! Front Page Today Module User Click Log Dataset. We observe a set of common characteristics of the data in the news domain[3]. They are:

- Average consumption time - The engagement time of a user with a news article is significantly low compared to other interactions like movies,musics, books or products. This means that user's short-term preferences constantly evolve with time.
- Lifespan of a News Article - As discussed in the data analysis section, the lifespan(survival time) of a typical news article is less than 2 days. The shelf-life of news articles are much shorter compared to music, movies, products.
- Number of Articles - News articles flood the internet every single day. In case of a trending event, thousands

of news articles are generated and circulated with minutes. The proliferation rate of news is extremely high compared to other items like music or movies who have relatively limited catalog size (several hundreds to few thousands).
- Sequential Consumption - Users tend to avoid news recommended sequentially if the articles are similar. Thus it becomes imperative to consider diversity in the domain while recommending news. Also from the standpoint of a society, its important to maintain balance and not creating an unwarranted bias by recommending the same content across groups.

These characteristics of the dataset make the task of news recommendation quite interesting and challenging to work on.

4.2 News Recommendation Literature

News recommendation has been a promising area of research over the last couple decades. The main challenge concerning this task was the lack of availability of a good benchmark dataset. Researchers who wished to work on the problem of news recommendation generally scraped their own datasets for a long time before switching to some datasets that were mentioned above.

The most common challenges faced by the news recommender models are: Cold-start, Data sparsity, Implicit user feedback, Scalability, Timeliness and Recency. State-of-the-art and traditional models employed creative and novel approaches to primarily tackle these challenges. Some of the past models are briefly described below:

- **Collaborative Filtering** - Memory-based collaborative filtering methods compared users among themselves using some similarity measures. The main issues observed with CF models were cold-start and data sparsity[1]. The number of news articles is much higher compared to the number of users in most systems which leads to data sparsity that in-turn leads to a degrade in the performance of CF based models. As mentioned earlier, news articles on news websites are posted continuously, get updated quite frequently and tend to expire very soon. This leads to cold-start problem being very severe CF models. While implementing similarity based models we experienced issues consistent with the problems described above. CF based methods also do not scale well and so cannot be employed for real-time news recommendation.
- **Time Decay** - In several papers[10], "time decay" was employed to factor the "decay" in user preference with time. One such approach that we tried was to redefine the item-item similarity by introducing time decay. We tried employing convex time decay, exponential time decay patterns to our similarity-based recommendation models but did not see promising results. The reason for this is MIND dataset implicitly handles temporal dynamics as the train set is spaced out only over a week. For other news datasets, time decay methods may have worked definitely better owing to the timeliness and short survival span of news data.
- **Deep Learning based models** - Recent works involve deep neural networks to learn the underlying feature representations from news articles. In[7], an ensemble model is proposed where news articles are represented using denoising autoencoder and users are modelled based on their click history using GRU. In[8], Wang et al. proposes CNN to learn news representations. We did not try deep learning models for this task. We propose them as future work.
- **Factorization Models** - Several factorization models like Matrix factorization (MF)[6], Non-Negative Matrix Factorization (NMF)[11], Tensor Factorization (TF), BPR have been employed for news recommendation. We have experimented quite a bit with factor models and observed that they offer promising performance which is consistent with findings in literature review.

5 Results and Conclusion

The experimental results of various methods are summarized in Table 3. All of the attempted news recommendation approaches perform significantly better than the baselines (Table 4) except Jaccard similarity which under performs in comparison to popular category baseline predictors. This is due to the skewed nature of the dataset with a high fraction of the dataset belonging to very few categories. This also explains why the popular category per user based filtering recommendation also performs better than the baselines.

The results show a significant improvement in AUC and BER with the FastFM factorization machine implementation. When run only on seen users and items, the model performs exceedingly well close to ideal predictors. However, this is not of much use in a practical scenario and hence cold start needs to be handled. On handling cold start, the AUC and BER fall relatively lower.

The news categories, sub-categories and length of news title features collaborated with user and item encodings performed the best for the news recommendation problem. Overall, the results show the importance of accurate news content understanding and modeling user-article interactions for news recommendation.

The fastFM model can only take in handcrafted features as input along with learning interaction data. In the future, we can extend this model to DeepFM [8] - using deep neural

Table 3. Results on MIND test set

Model	AUC	BER
Jaccard Sim.	0.52	0.49
Popular news cat./user	0.573	0.426
BPR	0.54	0.48
Logistic Regression	0.61	0.41
FastFM (Seen only)	0.97	0.04
FastFM (Overall)	0.662	0.338

Table 4. Results of Baseline predictors

Model	AUC	BER
Zero predictor	0.5	0.5
Most popular category	0.531	0.472
Most popular category per day	0.53	0.475

networks with interaction data to uncover and learn non-linear features with minimal handcrafted feature extraction.

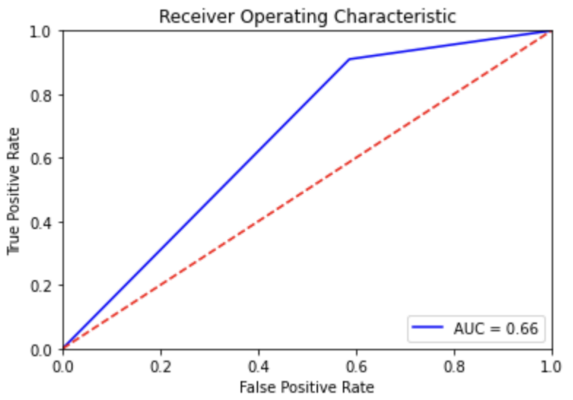


Figure 11. ROC Curve for FastFM on MIND test set

References

[1] Gediminas Adomavicius and Alexander Tuzhilin. 2005. Toward the next generation of recommender systems: A survey of the state-of-the-art and possible extensions. *IEEE transactions on knowledge and data engineering* 17, 6 (2005), 734–749.

[2] Immanuel Bayer. 2016. fastfm: A library for factorization machines. *The Journal of Machine Learning Research* 17, 1 (2016), 6393–6397.

[3] Abhijnan Chakraborty, Bhargavi Paranjape, Sourya Kakarla, and Niloy Ganguly. 2016. Stop clickbait: Detecting and preventing clickbaits in online news media. In *2016 IEEE/ACM international conference on advances in social networks analysis and mining (ASONAM)*. IEEE, 9–16.

[4] Jon Atle Gulla, Lemei Zhang, Peng Liu, Özlem Özgöbek, and Xiaomeng Su. 2017. The adressa dataset for news recommendation. In *Proceedings of the international conference on web intelligence*. 1042–1048.

[5] Benjamin Kille, Frank Hopfgartner, Torben Brodt, and Tobias Heintz. 2013. The plista dataset. In *Proceedings of the 2013 international news recommender systems workshop and challenge*. 16–23.

[6] Yehuda Koren, Robert Bell, and Chris Volinsky. 2009. Matrix factorization techniques for recommender systems. *Computer* 42, 8 (2009), 30–37.

[7] Shumpei Okura, Yukihiro Tagami, Shingo Ono, and Akira Tajima. 2017. Embedding-based news recommendation for millions of users. In *Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. 1933–1942.

[8] Hongwei Wang, Fuzheng Zhang, Xing Xie, and Minyi Guo. 2018. DKN: Deep knowledge-aware network for news recommendation. In *Proceedings of the 2018 world wide web conference*. 1835–1844.

[9] Fangzhao Wu, Ying Qiao, Jiun-Hung Chen, Chuhan Wu, Tao Qi, Jianxun Lian, Danyang Liu, Xing Xie, Jianfeng Gao, Winnie Wu, et al. 2020. Mind: A large-scale dataset for news recommendation. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*. 3597–3606.

[10] Chaolun Xia, Xiaohong Jiang, Sen Liu, Zhaobo Luo, and Zhang Yu. 2010. Dynamic item-based recommendation algorithm with time decay. In *2010 Sixth International Conference on Natural Computation*, Vol. 1. IEEE, 242–247.

[11] Jing Yang, Jing Wan, Yunxiang Wang, and Yan Mao. 2020. Social network-based News Recommendation with Knowledge Graph. In *2020 IEEE International Conference on Information Technology, Big Data and Artificial Intelligence (ICIBA)*, Vol. 1. IEEE, 1255–1260.