Enhancing Indonesian News Recommendations through Metadata Integration with Neural Attentive Multi-View Learning

Maxalmina Satria Kahfi, Evi Yulianti, Alfan Farizki Wicaksono

Faculty of Computer Science, Universitas Indonesia, Depok, Indonesia

Correspondence: Evi Yulianti (evi.y@cs.ui.ac.id)

ARTICLE INFO

Article history:
Received
Received in revised form
Accepted
Available online

Keywords:
News Recommendation; Named Entity Recognition; Recommendation Systems; User Modeling

ABSTRACT

The news recommendation system has the potential to help users discover articles that match their interests, which is crucial to alleviate user information overload. To generate effective news recommendations, one key capability is to accurately capture the contextual meaning of the text in news articles, as this is essential for obtaining useful representations for both news content and users. In this study, we examine the effectiveness of neural news recommendation with the Neural News Recommendation with Attentive Multi-View Learning (NAML) method to perform the news recommendation task in the Indonesian language. We investigate techniques for generating suitable vector representations of named entities in news content. We also propose to incorporate news metadata such as tags and entities in the news to improve the effectiveness of the NAML method in the Indonesian news recommendation system. Our results show that the NAML method leads to significant improvement in the effectiveness of news recommendations in the Indonesian language. Further addition of news metadata has been shown to improve the performance of the NAML method up to by 5.86% in terms of the NDCG@5 metric.

1. Introduction

Recommender systems are a central topic in information retrieval research, essential for helping users discover items aligned with their preferences [1]. In business, they are key to marketing strategies, particularly in boosting online sales through personalized recommendations [1,2]. Recommendation tasks are typically automated using content-based and collaborative filtering techniques [3], with recent advances leveraging embedding representations of users and items to predict user-item interactions more effectively. Recommender systems have been applied across diverse domains, including movie recommendation [4,5], news recommendation [6–12], and music recommendation [13,14].

News recommendation has become one of the most frequent applications of recommender systems in the digital landscape, especially for online platforms and news publishers [2,9,15]. Given the large volume of news articles generated daily, it is impractical for users to manually sift through all available content to find news that aligns with their interests [16]. The continuously increasing amount of news content published online presents a challenge for users in efficiently identifying articles that match their preferences [17]. Therefore, the implementation of personalized news recommendation systems has become essential, as it enables online news platforms to effectively target user preferences and mitigate the issue of information overload [9].

In recent years, English-language news recommendation models have demonstrated significant performance improvements through the application of deep learning techniques [18]. These models are commonly referred to as neural news recommendation models. Various neural architectures have been introduced to address challenges in news recommendation systems, one of which is Neural News Recommendation with Attentive Multi-View Learning (NAML) [11]. The NAML framework leverages an attentive multi-view learning approach to learn unified news representations from multiple types of information, such as news titles, content, and categories. This approach has been shown to outperform several existing deep learning methods for news recommendation, including Convolutional Neural Networks (CNNs) [19], Deep Fusion Models (DFM) [20], and Knowledge-aware CNN-based models such as DKN [6]. NAML also provides flexibility in incorporating additional auxiliary information, enabling the construction of more accurate news representations and subsequently producing more effective recommendations.

Prior research on English-language news recommendation systems has generally relied on a limited set of components, such as headlines, categories, and subcategories [21]. This focus, however, neglects the utility of fine-grained metadata. News tags, for example, contain latent semantic cues that can improve recommendation relevance [22]. In addition, named entities at various levels may capture key points within an article, thus offering supplementary information that contributes to better content representation [6]. This study addresses two primary research questions as follows:

I. What methodologies can be employed to generate effective entity vector representations for seamless integration and utilization within the NAML architecture?
II. What is the impact of various integration techniques for news metadata (e.g., tags and entities) on the efficacy of Indonesian news recommendation systems?

The structures of the paper are arranged as follows: we first review prior work in news recommendation in Section 2. In Section 3, we describe our research methodology and experimental design. We then present and analyze the results of our study in Section 4. Finally, in Section 5, we conclude with a summary of our contributions and provides recommendations for future work.

2. Literature Review

2.1. News Information

In news recommendation systems, effective representation of news content is essential for building accurate and personalized recommendation models [2,16]. These systems learn user preferences by modeling previously read articles, where components such as the title, abstract, body, and category are encoded into semantic vectors [21]. These representations capture the underlying meaning of news content, supporting user modeling and relevance-based recommendations [18,23]. Beyond standard text features, elements like news tags and named entities provide additional semantic cues. Tags—derived via social annotation or statistical techniques (e.g., TF-IDF, correlation metrics) [24,25]—highlight key topics and have been shown to improve performance in Indonesian language [22]. Named entities (e.g., people, organizations, locations) act as implicit signals, and recent models like HieRec leverage hierarchical or self-attention over both textual and entity embeddings to enhance contextual modeling [26].

2.2. Text Embedding

Text embedding is a technique that represents textual information as real-valued vectors. This allows computers to process text in a more meaningful way, as vector representations can capture not only syntactic but also semantic information. Several techniques have been used to do text embedding such as Word2Vec [27], Fasttext [28], GloVe [29], and BERT [30]. The use of text embeddings has demonstrated strong performance across a variety of tasks, including text summarization [31,32], ranking in expert search tasks [33,34], and text classification [34–36]. In the context of this research, text embedding will be applied to textual data found in news information, particularly in elements such as titles, abstract, content, and tags.

2.3. Entity Embedding

An entity needs to be transformed into a vector representation in order to be processed by a news recommendation system. In recent developments, entity representations can be constructed using a knowledge graph [21]. A knowledge graph is typically derived from a knowledge base and is often represented as knowledge triples that describe entities and their relational metadata. These triples follow the format (head, relation, tail), where both the head and the tail are entities, and the relation denotes the relationship between them. These relationships collectively represent the semantic structure within the knowledge graph. Based on these knowledge triples, entities can be modeled to extract their vector representations. One of the widely used methods to learn such representations is TransE [37]. The core idea of TransE is to model the relationship between entities as a translation in the embedding space. More specifically, the model learns to approximate using Eq. (1).

$$e_{head} + e_{relation} \approx e_{tail} \tag{1}$$

This formulation is then refined and normalized using the L_p norm—typically the L_1 (Manhattan) or L_2 (Euclidean) norm. In many optimization tasks and loss functions, the goal is to maximize the score (the higher the better), rather than to minimize the distance. Thus, to convert the distance minimization into a score maximization problem, a negative sign is applied, resulting in the TransE interaction function in Eq. (2).

$$f(h, r, t) = -\|e_{head} + e_{relation} - e_{tail}\|_p \tag{2}$$

The objective is to minimize the value of $f(h, r, t)$ for correct entity triples, so that the resulting embeddings effectively capture the relational structure in the knowledge graph. This relational understanding can subsequently enhance the performance of the news recommendation system.

2.4. News Recommendation Methods

With the advancement of deep learning techniques, several studies have explored the application of neural networks in the field of news recommendation, commonly referred to as neural news recommendation [18,23]. There are two main categories of user behavior modeling approaches in neural news recommendation systems: Candidate-Agnostic (C-AG) and Candidate-Aware (C-AW) models [38]. In Candidate-Agnostic (C-AG) models, the user encoder constructs the user embedding solely based on the embeddings of previously clicked news articles, without incorporating the information from the candidate news articles [7,10–12]. As a result, the user embedding remains the same regardless of which candidate news articles are being evaluated. In contrast, Candidate-Aware (C-AW) models generate user embeddings that are influenced by the content of candidate news articles. This means the user representation can vary depending on the candidate news article being considered, allowing for more adaptive and context-aware recommendations [6,8].

One of the C-AG model is NAML [11], uses additive attention to encode user preferences. This mechanism captures and emphasizes key aspects of user interests, improving recommendation performance. For the purposes of this research, particular emphasis will be placed on the NAML model, given its strong performance and flexibility—especially in integrating various types of news content into the model. Its architecture is easily modifiable to incorporate additional news-related information. In user modeling, NAML employs an attention network to identify important clicked articles, enabling the extraction of a more informative user representation. Experimental results have demonstrated that deep learning approaches, such as NAML, yield significantly better performance compared to traditional machine learning-based recommendation methods. As such, NAML is selected as the primary model for this research.

Although several studies have addressed the problem of news recommendation in the Indonesian language, none to date have adopted deep learning-based or neural news recommendation approaches. One such study applied association rule mining, followed by computing similarity between article titles to produce a ranking score for recommendation [39]. Another study employed TF-IDF to obtain vector representations of article titles and utilized K-Nearest Neighbors (KNN) with cosine similarity for ranking [40].

3. Methodology

3.1. Dataset Collection

3.1.1. User Behavior

In this study, the news recommendation dataset was sourced from one of Indonesia's largest news portals and differs from that used in Kahfi *et al.* [22], particularly in its time span. The dataset, which is proprietary and not publicly available, includes user interactions from March 4 to April 7, 2024. Users were randomly selected based on having clicked at least five news articles during this five-week period, ensuring sufficient behavioral data for user modeling. The selected timeframe was intentionally chosen to avoid periods dominated by political events, which could bias the distribution of news categories. All user interactions were organic and uninfluenced by existing recommendation systems. The data collection process is illustrated in Figure 1.

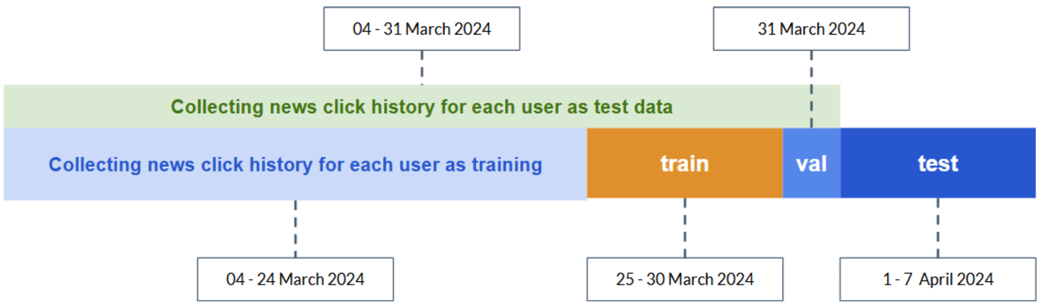

Figure 1. The workflow and time period of dataset collection are centered on user activity, which serves as the basis for the dataset used in this study

The dataset in this study is divided into three subsets: training, validation, and test. Each subset contains two types of data: user click history and session-based user activity. Since news recommendation involves predicting user clicks on candidate articles based on inferred preferences, click history was embedded into impression logs to generate labeled data for model training and evaluation. The training set and validation set use click data from the first three weeks to construct user profiles. Session-based user activity from the fourth week was used for evaluation, with the final day of that week reserved for validation. The test set includes click history from the first four weeks—containing unseen behavior—and is evaluated using session data from the fifth week. All articles published and interacted with during the five-week period were collected to support model input.

The number of samples in the training, validation, and test sets are 133,679, 31,377, and 24,160, respectively. User activities were converted into impression logs, which record viewed and clicked articles along with timestamps and interaction metadata. Dataset statistics are detailed in Table 1.

Table 1. Statistical information of the data constructed and utilized in this study

No	Component	Value
1	Number of Users	58,326
2	Number of User Logs	189,216
3	Number of News Articles	26,660
4	Number of Categories	18
5	Number of Subcategories	75
6	Number of Entities	-
7	Avg. Number of Tags per Article	4
8	Avg. News Title Length	10.8
9	Avg. News Abstract Length	17.3
10	Avg. News Content Length	351.3

As each impression only records the articles clicked by the user, this study employs negative sampling, in which four unclicked articles are sampled for every one clicked article in the impression. These negative samples are selected from the same day on which the impression occurred. This strategy helps address the label incompleteness issue commonly found in recommendation systems by ensuring that both positive and negative instances are considered during training.

3.1.2. News Entity

In this study, due to the limited availability of entity-related data in our Indonesian news articles dataset, several pre-trained language models (PLMs) will be evaluated and fine-tuned that conducted in separate experiment. The extracted named entities from the NER model will be transformed into vector representations, and two different approaches are explored for constructing these vectors. The effectiveness of each representation will be assessed based on the overall performance of the news recommendation system.

The first approach utilizes both the entity label (i.e., the text span) and its corresponding entity category as identified by the NER model. The entity label is converted into a vector using a pre-trained word embedding model. Meanwhile, the entity type is encoded using one-hot encoding. The resulting vectors from the entity label and type are concatenated to form a single representation for the entity. To facilitate computational efficiency, the resulting entity vectors are then reduced to a fixed number of dimensions—100 or 300—using dimensionality reduction techniques.

The second approach involves entity linking, where the identified entities are linked to a knowledge base, specifically Wikidata. Through entity linking, each entity is associated with a unique Wikidata ID, enabling the retrieval of corresponding knowledge triples. These triples are instrumental in constructing knowledge-based embeddings of entities. Moreover, entity linking allows for entity disambiguation, ensuring that entities with multiple aliases are represented by a single, unified vector. Once the knowledge triples are retrieved, they are used to train entity embeddings using the TransE algorithm, which learns vector representations from the graph structure. TransE effectively considers both incoming and outgoing edges of the entities during training, resulting in rich and semantically meaningful embeddings.

3.2. News Recommendation Model

In this study, NAML is utilized as the primary model. The architecture is generally divided into several major components, each of which plays a specific role in the news recommendation process. For the purpose of this research, the original NAML architecture is modified to incorporate various types of information from different news components. The illustration of the modified NAML news encoder architecture is shown in Figure 2.

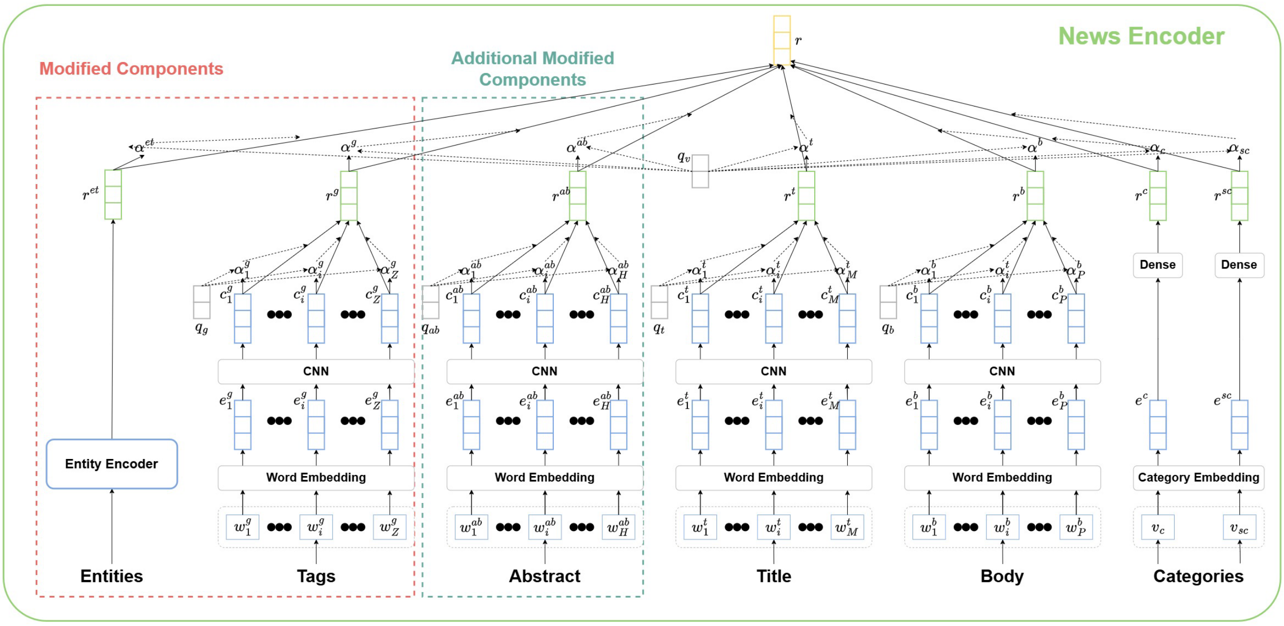

Figure 2. Architecture of the news encoder in the modified NAML model for processing different types of news information

The core component of the NAML architecture is the news encoder, which transforms the input—i.e., various elements of a news article such as its category, title, and content—into a rich vector representation suitable for machine learning models. We modified the basic NAML architecture to enable the model to utilize all news metadata such as tags and entities in the news. To process information components within a news article, several subcomponents are used.

3.2.1. Text Encoder

The first subcomponent is the text encoder, which processes the textual content found in the title, abstract, body, and tags of the news article. The first layer of this subcomponent is the word embedding layer, which transforms the sequence of words in the news title into a sequence of low-dimensional semantic vectors. These vectors capture the semantic meanings of the words. Formally, the word sequence in a news title is denoted as $[w_1^t, w_2^t, \ldots, w_M^t]$, where M is the length of the title. This sequence is transformed into word embeddings $[e_1^t, e_2^t, \ldots, e_M^t]$ using a word embedding lookup table $W_e \in \mathcal{R}^{V \times D}$, where V is the vocabulary size and D is the embedding dimension.

The second layer is a Convolutional Neural Network (CNN), which extracts important local features from the word embeddings. CNNs are effective in capturing local context within the text. If the contextual representation of the i-th word is denoted as c_i^t, then it is computed using Eq. (3).

$$c_i^t = ReLU(F_t \times e_{(i-K):(i+K)}^t + b_t) \tag{3}$$

Here, $e^t_{(i-K):(i+K)}$ represents the concatenated embeddings from position $(i-K)$ to $(i+K)$; $F_t \in \mathcal{R}^{N_f \times (2K+1)D}$ is the filter matrix; and $b_t \in \mathcal{R}^{N_f}$ is the bias vector. N_f denotes the number of CNN filters and $2K+1$ is the convolution window size. The output is a contextual representation sequence $[c^t_1, c^t_2, \dots, c^t_M]$. The final layer is a word-level attention network, which weighs the importance of each word within the textual component of the news. This layer allows the model to identify which words contribute more to the overall semantics of the text. For example, in the headline "*Liverpool Mengakhiri Tahun 2019 dengan Sebuah Kemenangan*" (Liverpool Ends 2019 with a Victory), the word "Liverpool" may be assigned a higher weight as it likely represents the subject of the news. The attention score α^t_i for the i-th word is computed using Eq. (4).

$$a^t_i = q^T_t \, tanh(V_t \times c^t_i + v_t),$$
$$\alpha^t_i = \frac{exp(a^t_i)}{\sum_{j=1}^{M} exp(a^t_j)} \quad (4)$$

Where V_t and v_t are projection parameters, and q_t is the attention query vector. The final representation of the news title, r^t, is constructed by aggregating its contextual word representations through a weighted summation process, as shown in Eq. (5). The remaining news information types—abstract (r^{ab}), content (r^b), and tags (r^g)—are computed using the same procedure employed for the title vector representation.

$$\mathrm{r}^t = \sum_{j=1}^{M} \alpha^t_j c^t_j \quad (5)$$

3.2.2. Entity Encoder

The entity encoder is another subcomponent, designed to process entity information present in news articles. In this study, two experimental scenarios were conducted using different entity encoding methods. These scenarios are illustrated in Figure 3.

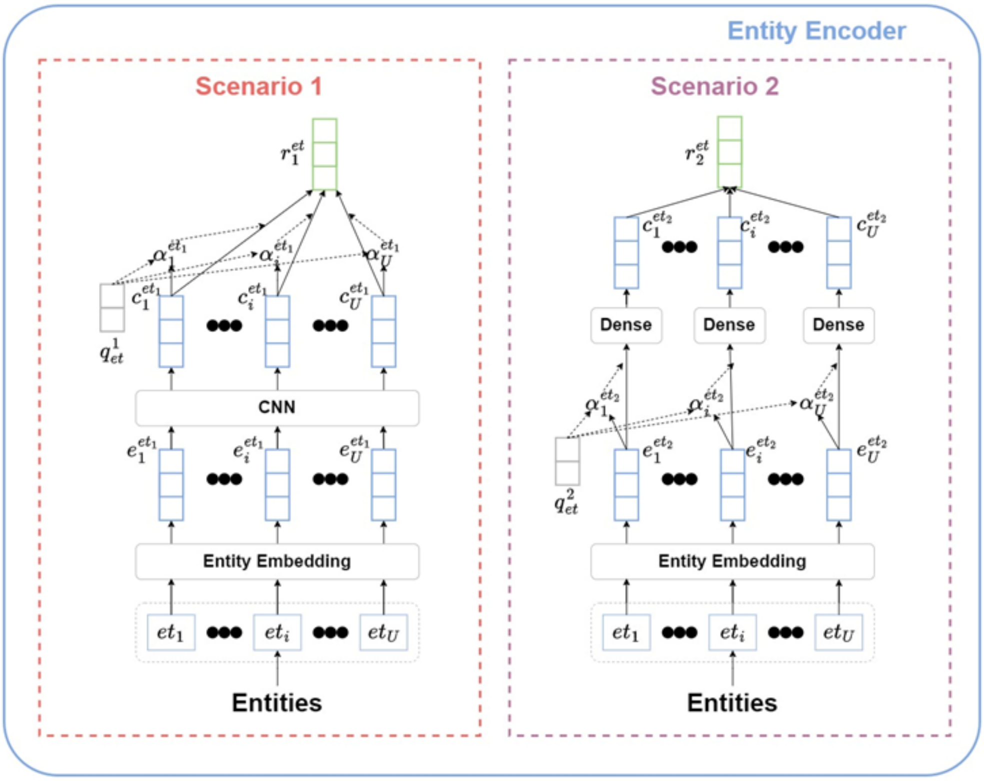

Figure 3. Two experimental scenarios for the entity encoder in NAML

In the first scenario, the entity encoder consists of three layers, similar to the preceding three subcomponents. The first layer is the entity embedding layer. Each entity in a news article is formally represented as $[et_1, et_2, \ldots, et_U]$, where U denotes the number of entities in the article. Each entity is then transformed into a vector representation $\left[e_1^{et_1}, e_2^{et_1}, \ldots, e_U^{et_1}\right]$ using an entity embedding lookup table $Et_e \in \mathcal{R}^{W \times B}$, where W is the total number of entities and B is the embedding dimension.

The second layer is a Convolutional Neural Network (CNN) layer, which generates contextual representations of the entities, denoted as $[c_1^{et_1}, c_2^{et_1}, \ldots, c_U^{et_1}]$, using the output from the previous embedding layer as input. This helps the model identify which representation best captures the semantics of each entity. The final layer is an attention layer, used to construct the final entity representation in the article. This layer assigns higher weights to more informative entities, enabling the model to emphasize semantically rich entities over less relevant ones. The attention weight for the i-th entity is denoted by $\alpha_i^{et_1}$, and is calculated using Eq. (6).

$$a_i^{et_1} = q_{et_1}^T \, tanh\left(V_{et_1} \times c_i^{et_1} + v_{et_1}\right),$$
$$\alpha_i^{et_1} = \frac{exp(a_i^{et_1})}{\sum_{j=1}^{U} exp(a_j^{et_1})} \tag{6}$$

Here, V_{et_1} and v_{et_1} are projection parameters, and q_{et_1} is the attention query vector. The final entity representation in scenario 1 is computed using Eq. (7).

$$r_1^{et} = \sum_{j=1}^{U} \alpha_j^{et_1} c_j^{et_1} \tag{7}$$

In the second scenario, the entity encoder also comprises three layers. Similar to Scenario 1, the first layer uses an entity embedding layer to convert each entity $[et_1, et_2, \ldots, et_U]$ into a vector representation $[e_1^{et_2}, e_2^{et_2}, \ldots, e_U^{et_2}]$. In the second layer, these embeddings are passed through an attention layer to obtain a weighted vector representation that emphasizes more important entities within the article. The i-th entity's attention weight $\alpha_i^{et_2}$ is computed using Eq. (8).

$$a_i^{et_2} = q_{et_2}^T \tanh\left(V_{et_2} \times c_i^{et_2} + v_{et_2}\right),$$
$$\alpha_i^{et_2} = \frac{exp(a_i^{et_2})}{\sum_{j=1}^{U} exp(a_j^{et_2})} \tag{8}$$

Where V_{et_2} and v_{et_2} are projection parameters, and q_{et_2} is the attention query vector. The resulting representation from the second layer is calculated using Eq. (9).

$$c_i^{et_2} = \sum_{j=1}^{U} \alpha_j^{et_2} e_j^{et_2} \tag{9}$$

The final layer is a dense layer, parameterized by V_l and v_l (as in Eq. (10)), which produces the final entity representation. This layer enables the model to identify and exploit more abstract and complex relationships among entities within the article.

$$r_2^{et} = ReLU(V_l \times c_i^{et_2} + v_l) \tag{10}$$

3.2.3. Categorical Encoder

In contrast to the processing of textual components, the next subcomponent of the architecture deals with categorical metadata, specifically the category and subcategory of the news articles. This information is processed using a category encoder, which consists of two neural network layers. The first layer serves to transform the discrete category ID and subcategory ID into low-dimensional dense vector representations, denoted as e^c and e^{sc}, respectively. These embeddings reduce the sparsity and complexity of the original categorical inputs, enabling the model to better learn meaningful patterns and semantic relationships.

The second layer is a fully connected (dense) layer that refines the embeddings learned in the first layer. This layer is responsible for learning deeper and more abstract representations of the category and subcategory by transforming the initial embeddings into higher-level features. Through this transformation, the model can capture and leverage more nuanced relationships between categories and subcategories—relationships that may not be immediately evident from the raw IDs alone. The final representations of the category and subcategory are calculated using the Eq. (11).

$$r^c = ReLU(V_c \times e^c + v_c)$$
$$r^{sc} = ReLU(V_{sc} \times e^{sc} + v_{sc}) \tag{11}$$

3.2.4. Attentive Pooling

The final subcomponent of the news encoder is the attentive pooling module. Different types of news information may carry varying degrees of informativeness for learning comprehensive news representations. For instance, a news title with a greater number of words and high informational content should contribute a larger weight in representing the article. Conversely, if the title is short and contains limited information, other components—such as the body/content, category, or additional metadata—should be assigned higher weights in representing the news article.

To address this variability, a view-level attention network is employed to model the informativeness of each type of news information, thereby enhancing the learning of a more effective news representation. The attention weights corresponding to each view—title, body, abstract, tags, category, subcategory, and entities—are denoted as α_t, α_b, α_{ab}, α_g, α_c, α_{sc}, α_{et}, respectively. The attention weight for the title view is computed using Eq. (12). The attention weights for the other types of news information are calculated in a similar manner, following the same procedure as that used for the title view.

$$a_t = q_v^T \, tanh(U_v \times r^t + u_v),$$
$$\alpha_t = \frac{exp(a_t)}{exp(a_t) + exp(a_b) + exp(a_{ab}) + exp\left(a_g\right) + exp(a_c) + exp(a_{sc}) + exp(a_{et})} \tag{12}$$

Finally, all subcomponent representations are aggregated to form a unified output that encapsulates the full representation of a news article. The final news representation vector is constructed by computing the weighted sum of the representations from each view, modulated by their respective attention weights. This aggregation is formalized in Eq. (13).

$$r = \alpha_t r_t + \alpha_{ab} r_{ab} + \alpha_b r_b + \alpha_c r_c + \alpha_{sc} r_{sc} + \alpha_g r_g + \alpha_{et} r_{et} \tag{13}$$

3.2.5. User Encoder

Another critical component is the user encoder, which transforms a user's historical interactions with previously read news articles into a numerical representation that models the user's preferences. The user encoder employs an attention network, a computational model capable of focusing on the most relevant parts of the input data—in this case, the news articles that the user has consumed. For example, a news article titled *“5 momen terbaik Persija di Liga 1”* (5 Best Moments of Persija in Liga 1) may be highly informative in modeling a particular user if the user has a consistent history of reading sports-related content. In contrast, an article titled *“Jadwal KRL Bekasi Line”* (Bekasi Line Commuter Schedule) may be less informative, as it is generic and likely to be clicked by a broad and diverse user base. Therefore, the user encoder is designed to extract informative signals from the user’s reading history by leveraging the numerical representations generated by the news encoder for each clicked article. The attention weight for the i-th news article clicked by the user is denoted as α_i^n, which is computed using Eq. (14).

$$a_i^n = q_n^T \, tanh(W_n \times r_i + b_n),$$
$$\alpha_i^n = \frac{exp(a_i^n)}{\sum_{j=1}^{N} exp(a_j^n)} \tag{14}$$

Here, q_n, W_n, and b_n are learnable parameters of the news attention network. The output of the user encoder is the weighted sum of the representation vectors of all news articles the user has interacted with, computed according to Eq. (15).

$$u = \sum_{i=1}^{N} \alpha_i^n r_i \tag{15}$$

3.2.6. Click Predictor

For The final component is the click predictor, which computes the likelihood that a user will click on a candidate news article. This component takes as input the representation of the candidate news article D^c, which has been encoded by the news encoder into a vector r_c, and the user's preference representation u, as computed by the user encoder. The click probability $\hat{y}$ is then calculated using the inner product between the user vector u and the candidate news vector r_c as formatted in the Eq. (16).

$$\hat{y} = u^T r_c \tag{16}$$

This probability measures how likely it is that a user will be interested in and click on the recommended news article. The use of the inner product not only provides an effective measure of relevance between user and news representations but also offers computational efficiency. This makes it particularly well-suited for large-scale recommendation systems, as demonstrated in prior research [17].

3.3. Experiment Settings

For To train the NAML model, several hyperparameters were utilized. All components and corresponding hyperparameter values are summarized in Table 2. The news recommendation system was trained using Google Colaboratory with a GPU A100 hardware accelerator.

Table 2. Hyperparameter settings used in the model

No	**Parameter**	**Value**
1	Word Embedding Dimension	300
2	Category Embedding Dimension	100
3	Entity Embedding Dimension	[100, 300]
4	News Vector Dimension	400
5	User Vector Dimension	400
6	Epochs	2
7	Batch Size	32
8	Negative Sampling Ratio	2
9	Dropout Probability	0.2
10	Query Vector Dimension	200
11	Number of Filters	300
12	Window Size	3

To evaluate the performance of the News Recommendation System (NRS), this study employed several evaluation metrics encompassing both accuracy (AUC) and relevance (MRR and NDCG), along with hypothesis testing. These three metrics are widely used as standard benchmarks in neural news recommendation evaluations [18,23,38]. In terms of model training, cross-entropy loss with negative sampling is applied as the optimization objective. This loss function serves as a straightforward classification objective for training the news recommendation model [41]. Moreover, hypothesis testing is conducted to assess the statistical significance of performance differences among recommendation models. This is done using Tukey's Honestly Significant Difference (HSD) test, a statistical method designed for multiple comparison analysis. If the mean difference between any two treatments exceeds this HSD threshold, the difference is deemed statistically significant. Otherwise, no significant difference can be inferred.

4. Results

4.1. Preliminary Results

The results of observations on various neural network-based news recommendation methods for Indonesian-language news are presented in Table 3. These results demonstrate that different news recommendation models exhibit diverse performance levels across a range of evaluation scenarios. The evaluations reveal that the primary model under observation performs well in terms of both accuracy and relevance. Furthermore, from a computational efficiency perspective (i.e., model training time), NAML stands out as one of the most efficient methods compared to others. Compared to the baseline approach implemented in the news publisher's existing news recommendation system (KCM), NAML achieves significant improvements of 11.25%, 10.76%, 11.25%, and 9.55% in AUC, MRR, NDCG@5, and NDCG@10, respectively. This makes NAML the most suitable method for addressing news recommendation challenges in the Indonesian language.

Table 3. Performance of Various News Recommendation Models. Significant differences compared to baseline models LSTUR/TANR/NRMS/CAUM are indicated using † / ‡ / ● / ∗ / ♦ / ◊ for $p < 0.001$. Computational time is represented in HH:MM:SS format

Method	AUC	MRR	NDCG@5	NDCG@10	Running Time
LSTUR-ini†	0.5913	0.4454	0.5348	0.5926	00:07:43
LSTUR-con‡	0.6351	0.4504	0.5500	0.6008	00:07:47
KCM●	0.7145	0.5575	0.6483	0.6870	00.03.27
TANR∗	0.7974	0.6307	0.7277	0.7535	00:06:51
NRMS♦	0.8016	0.6411	0.7378	0.7623	00:05:03
CAUM◊	0.8077	0.6462	0.7438	0.7669	00:12:22
NAML	**0.8282†‡●∗♦◊**	**0.6651†‡●∗♦◊**	**0.7608†‡●∗♦◊**	**0.7825†‡●∗♦◊**	**00:05:06**

4.2. Entity Embedding Performance

Prior to the integration of news entities into the model, two distinct workflows were investigated for the generation of entity embeddings and their application within a news recommendation system. The two entity embedding models were then implemented via integration into the news recommendation system. In this part of the study, we incorporated not only entity information but also basic article features, including the title, category, and subcategory. These components were retained to enable a performance comparison across different representation methods.

Empirical results, detailed in Table 4, demonstrated that higher dimensionality (i.e., 300 dimensions) achieved uplift performance. The entity linking to a knowledge base method outperformed the baseline by up to 1.53%, 2.89%, 2.23%, and 1.96% in terms of AUC, MRR, NDCG@5, and NDCG@10, respectively. Furthermore, the outcomes obtained with and without the application of entity linking did not differ substantially. This suggests that both embedding techniques yielded comparable performance with no statistically significant distinction. In a broader context and across multiple cases, the utilization of a CNN and an additive attention mechanism to process the vector representations of news entities proved more effective than relying solely on a linear layer.

Table 4. Performance metrics of recommendation systems before and after adding named entities. Significant differences are indicated using ‡ / ● / ∗ / ♦ for $p < 0.001$

Method	Dimension	Entity Encoder	AUC	MRR	NDCG@5	NDCG@10	Running Time
Baseline (No Entities Included)	-	-	0.8282	0.6651	0.7608	0.7825	5min 6s
Word Embedding + Entity Type One-Hot Encoding	100	Linear Layer	0.8376	0.6783	0.7735	0.7932	6min 16s
	100	CNN + Linear Layer	0.8325	0.6757	0.7701	0.7905	7min 23s
	300	Linear Layer	0.8206	0.6586	0.7543	0.7765	6min 34s
	300	CNN + Linear Layer	0.8396	0.6834	0.7775	0.7971	10min 33s
Entity Linking - TransE Embedding	100	Linear Layer	0.8328	0.6732	0.7680	0.7886	6min 13s
	100	CNN + Linear Layer	0.8309	0.6723	0.7673	0.7879	7min 22s
	300	Linear Layer	0.8302	0.6680	0.7634	0.7849	6min 24s
	300	**CNN + Linear Layer**	**0.8409**	**0.6843**	**0.7778**	**0.7978**	**10min 20s**

4.3. Integration of Named Entity

The utilization of news entities has a significant impact on enhancing the performance of a news recommendation system. As illustrated in Table 5, which presents a comparative analysis of the system's performance with and without entity integration, this effect is evident. The integration of entities into the recommendation model significantly boosted the performance of the NAML method, yielding improvements of up to 4.63% in AUC, 7.49% in MRR, 5.86% in nDCG@5, and 5.05% in nDCG@10. While this enhancement did not surpass the performance gains from using news tags, it nonetheless constitutes a significant improvement.

Table 5. Performance metrics of recommendation systems before and after adding named entities. Significant differences are indicated using ‡ / ● / ∗ / ♦ for $p < 0.001$

News Information Component	AUC	MRR	NDCG@5	NDCG@10
Title	0.8032	0.6368	0.7350	0.7591
Abstract	0.7765	0.6102	0.7075	0.7358
Content	0.7920	0.6209	0.7201	0.7459
Title + Category/Subcategory‡	0.8282	0.6651	0.7608	0.7825
Title + Abstract●	0.7947	0.6274	0.7253	0.7511
Title + Content	0.8168	0.6515	0.7498	0.7717
Title + Category/Subcategory + Abstract∗	0.8170	0.6576	0.7525	0.7755
Title + Category/Subcategory + Content	0.8362	0.6756	0.7718	0.7914
Title + Abstract + Content	0.8127	0.6482	0.7460	0.7688
Title + Category/Subcategory + Abstract + Content♦	0.8235	0.6612	0.758	0.7790
Title + Entity	0.8097	0.6418	0.7389	0.7628
Abstract + Entity	0.7696	0.5994	0.6978	0.7274
Content + Entity	0.7967	0.6341	0.7315	0.7566
Title + Category/Subcategory + Entity	**0.8415**‡	**0.6849**‡	**0.7787**‡	**0.7981**‡
Title + Abstract + Entity	**0.8315**●	**0.6744**●	**0.7678**●	**0.7890**●
Title + Content + Entity	0.8068	0.6405	0.7388	0.7625
Title + Category/Subcategory + Abstract + Entity	**0.8256**∗	**0.6610**∗	**0.7577**∗	**0.7788**∗
Title + Category/Subcategory + Content + Entity	0.8318	0.6687	0.7659	0.7852
Title + Abstract + Content + Entity	0.8024	0.6374	0.7354	0.7599
Title + Category/Subcategory + Abstract + Content + Entity	**0.8361**♦	**0.6762**♦	**0.7716**♦	**0.7916**♦

This performance gap is likely due to unresolved entity ambiguity, where entities that should possess distinct vector representations share the same label. For instance, an entity labeled "Megawati" necessitates contextual disambiguation to differentiate between "Megawati," the politician, and "Megawati," the volleyball athlete. The current lack of such a validation mechanism results in a single vector representation for these two distinct individuals. This ambiguity adversely affects the overall system performance, explaining why it does not match the efficacy of the model augmented with news tags.

Using the user's reading history (Figure 4), the news recommendation system generates suggestions as shown in Figure 5, which compares models with and without entity integration. Without entity information, relevant articles appeared lower in the ranking (e.g., third position). In contrast, incorporating news entities significantly improved recommendation quality, with the most relevant article moving to the top rank. Additional relevant articles also appeared within the top 5, indicating enhanced relevance and effectiveness through entity-aware modeling.

Respons Yusril Usai Pernyataan "MK Jangan Jadi Mahkamah Kalkulator" Dikutip Mahfud
Yusril's Response After His "MK Should Not Become a Calculator Court" Statement Was Quoted by Mahfud

Tags: *mahfud md, yusril ihza mahendra, yusril mk, mahkamah kalkulator*

Entities: [{'Label': ['Yusril','Mahfud'], 'Type': 'PER'}, {'Label': 'MK', 'Type': 'NOR'}]

Akademisi UGM Usul Pengadilan Rakyat, Moeldoko: Kita Negara Hukum, Jangan Diselesaikan dengan Cara Jalanan
UGM Academic Proposes People's Court, Moeldoko: We Are a State of Law, Don't Resolve Issues Through Street Methods

Tags: *ugm, moeldoko, kecurangan pemilu, pemilu 2024*

Entities: [{'Label': 'UGM', 'Type': 'ORG'}, {'Label': 'Moeldoko', 'Type': 'PER'}]

Sudah 1,5 Tahun Kompolnas dan Polisi Belum "Update" Kasus Kematian Akseyna
After 1.5 Years, National Police Commission and Police Have Not "Updated" Akseyna Death Case

Tags: *akseyna, akseyna ahad dori, akseyna kasusnya bukan bunuh diri, namun korban pembunuhan, akseyna ahad dori alias ace*

Entities: [{'Label': 'Polisi', 'Type': 'NOR'}, {'Label': 'Akseyna', 'Type': 'PER'}]

Surya Paloh: Kita Dipaksa pada Kepentingan Sesaat, Jangka Pendek yang Pragmatis
Surya Paloh: We Are Being Forced Into Momentary, Short-Term Pragmatic Interests

Tags: *nasdem, surya paloh*

Entities: [{'Label': 'Surya Paloh', 'Type': 'PER'}]

Respons Polri soal TPN Ganjar-Mahfud Akan Datangkan Kapolda di Sidang MK
Police Response Regarding TPN Ganjar-Mahfud Will Bring Regional Police Chiefs to MK Hearing

Tags: polri, ganjar-mahfud,mk

Entities: [{'Label': ['Polri','TPN', 'Kapolda'], 'Type': 'NOR'}, {'Label': 'Mahfud', 'Type': 'PER'}, {'Label': 'Sidang MK', 'Type': 'EVT'}]

[POPULER NASIONAL] Rakyat Dianggap Tidak Boleh Kalah dari Keluarga Jokowi | PSI Usul Jokowi Pimpin Koalisi
[NATIONAL POPULAR] People Considered Must Not Lose to Jokowi's Family | PSI Proposes Jokowi Lead Coalition

Tags: joko widodo, partai solidaritas indonesia (psi), jokowi, ikrar nusa bhakti

Entities: [{'Label': 'Jokowi', 'Type': 'PER'}, {'Label': 'PSI', 'Type': 'NOR'}]

Beda Sikap Anies dengan Nasdem-PKS Tanggapi Kemenangan Prabowo
Different Stance Between Anies and NasDem-PKS in Responding to Prabowo's Victory

Tags: anies baswedan, pemilu 2024, prabowo-gibran, anies-muhaimin, beda sikap anies dengan nasdem-pks

Entities: [{'Label': ['Anies','Prabowo'], 'Type': 'PER'}, {'Label': ['Nasdem','PKS'], 'Type': 'NOR'}

Golkar Minta Jatah 5 Menteri, Gerindra: Jangankan 5, kalau Kerjanya Maksimal Bisa Lebih
Golkar Requests 5 Ministerial Posts, Gerindra: Never Mind 5, If They Work Optimally They Could Get More

Tags: golkar, gerindra, menteri, prabowo-gibran

Entities: [{'Label': ['Golkar','Gerindra'], 'Type': 'NOR'}]

303 Guru Besar dan Akademisi Surati MK, Minta Hakim Tak Cuma Urusi Jumlah Suara Sengketa Pilpres
303 Professors and Academics Send Letter to MK, Ask Judges Not to Only Deal with Vote Count Disputes in Presidential Election

Tags: sengketa hasil pilpres 2024, sidang sengketa hasil pilpres, sidang sengketa pilpres 2024, sidang mk sengketa pilpres 2024

Entities: [{'Label': 'MK', 'Type': 'NOR'}, {'Label': 'Pilpres', 'Type': 'EVT'}]

Laporan terhadap Ganjar Dianggap Politisasi, Mahfud: Saya Tidak Pandang Itu, Terserah KPK Saja
Report Against Ganjar Considered Politicization, Mahfud: I Don't See It That Way, It's Up to the KPK

Tags: mahfud md, ganjar, ganjar dilaporkan ke kpk

Entities: [{'Label': ['Ganjar','Mahfud'], 'Type': 'PER'}, {'Label': 'KPK', 'Type': 'NOR'}

Figure 4. User's news reading history example

We found syntactic similarities between the user's news click history and the recommendation results when news entities were added. For example, "*Sidang*" (Trial) and "MK" (Constitutional Court) appeared in the click history, while "*Sidang* MK" and "MK" were present in recommended articles. Semantic similarity was also evident due to the frequent appearance of the entity "MK" in the user's news click history. This similarity improves the recommendation system's performance by incorporating news entity information. News entities in recommended articles tend to be more relevant to users. For instance, the entity "MK" appearing in both the user's history and the system's recommendations helps predict user preferences for that entity and provides relevant articles. Furthermore, even if the top-ranked relevant article doesn't have syntactic similarities in its title or entities, it remains relevant due to semantic similarities, such as a connection to state institutions (DPR in the recommendation and MK in the user's history).

Figure 5. Comparative results of the news recommendation system with and without news entity integration, based on a user's activity history

4.4. Integration of All News Metadata Results

The simultaneous integration of news tags and named entities yielded notable performance gains in the recommendation system. As shown in Table 6, incorporating both metadata types into the NAML model improved AUC by 2.41%, MRR by 3.04%, NDCG@5 by 2.68%, and NDCG@10 by 2.21%, compared to the baseline without both metadata. These enhancements confirm the value of leveraging both tags and entities. However, the combined integration did not outperform the individual incorporation of tags or entities (Table 5), suggesting that entity information contributed less to performance gains than tags. As a result, the joint use of both components did not produce a synergistic improvement.

Table 6. Performance metrics of recommendation systems before and after adding named entities. Significant differences are indicated using ‡ / ● / ∗ / ♦ for $p < 0.001$

News Information Component	AUC	MRR	NDCG@5	NDCG@10
Title‡	0.8032	0.6368	0.7350	0.7591
Abstract	0.7765	0.6102	0.7075	0.7358
Content	0.7920	0.6209	0.7201	0.7459
Title + Category/Subcategory	0.8282	0.6651	0.7608	0.7825
Title + Abstract●	0.7947	0.6274	0.7253	0.7511
Title + Content	0.8168	0.6515	0.7498	0.7717
Title + Category/Subcategory + Abstract∗	0.8170	0.6576	0.7525	0.7755
Title + Category/Subcategory + Content	0.8362	0.6756	0.7718	0.7914
Title + Abstract + Content	0.8127	0.6482	0.7460	0.7688
Title + Category/Subcategory + Abstract + Content♦	0.8235	0.6612	0.758	0.7790
Title + Tags + Entity	0.8119‡	0.6487‡	0.7455‡	0.7690‡
Abstract + Tags + Entity	0.7663	0.5946	0.6942	0.7236
Content + Tags + Entity	0.7775	0.6112	0.7096	0.7375
Title + Category/Subcategory + Tags + Entity	0.8348●	0.6763●	0.7712●	0.7916●
Title + Abstract + Tags + Entity	0.7931	0.6273	0.7247	0.7511
Title + Content + Tags + Entity	0.8043	0.6401	0.7378	0.7616
Title + Category/Subcategory + Abstract + Tags + Entity	0.8302∗	0.6721∗	0.7666∗	0.7876∗
Title + Category/Subcategory + Content + Tags + Entity	0.8370	0.6767	0.7729	0.7921
Title + Abstract + Content + Tags + Entity	0.8032	0.6385	0.7361	0.7604
Title + Category/Subcategory + Abstract + Content + Tags + Entity	0.8405♦	0.6806♦	0.7761♦	0.7952♦

We also conducted further analysis to understand how various combinations of news information impact the performance of the recommendation system. This analysis used NDCG@5 to measure the relevance of the recommendations and model training time to reflect computational efficiency. We also categorized the combinations of news information used. The results, shown in Figure 6, indicate that while adding both news tags and entities can improve recommendation performance in some cases, it generally leads to a significant increase in computational time.

For instance, adding only one additional information component (either news tags or news entities) doesn't cause as large an increase in computational time as adding both simultaneously. However, the performance improvement from adding both components isn't yet proportional to the increased computational time required. Therefore, while integrating more information can enhance recommendation quality, there's a trade-off between performance and computational efficiency that needs consideration. Further optimization may be necessary to find the best balance between these two aspects.

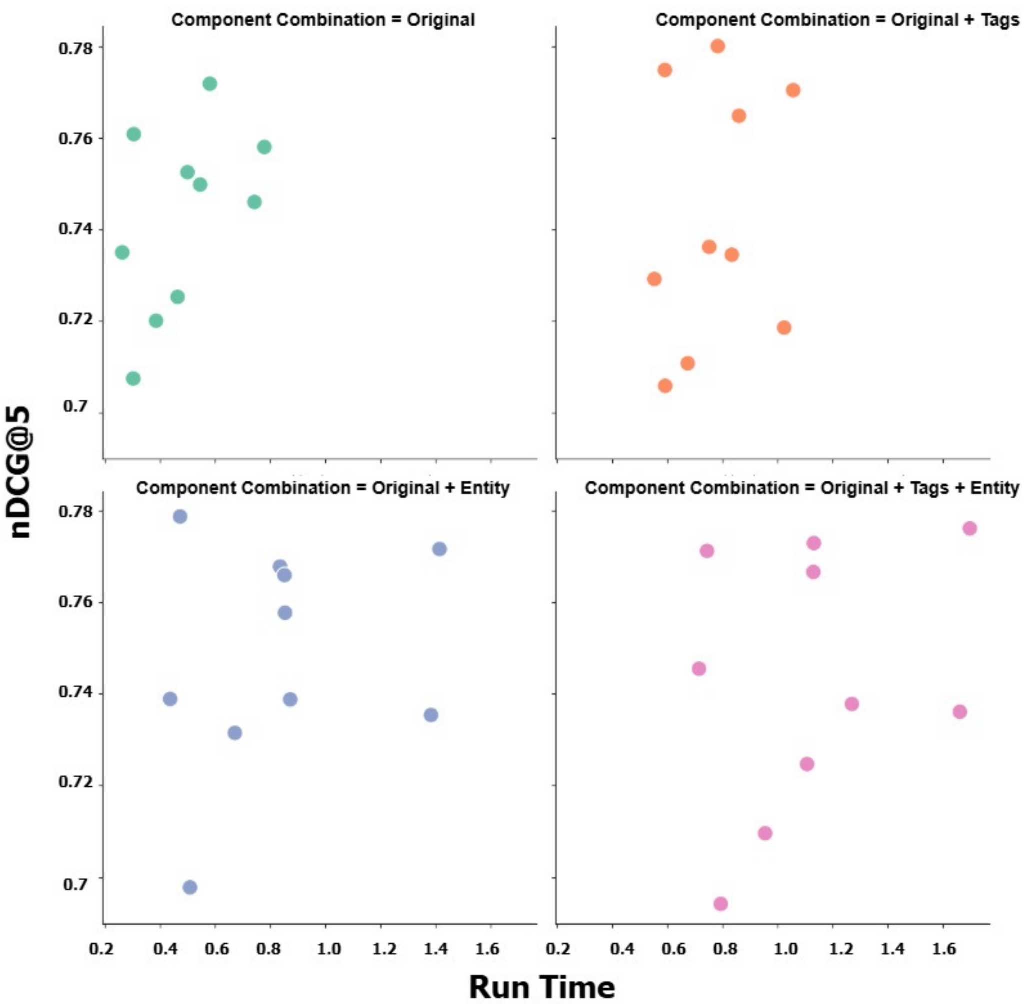

Figure 6. Analysis of the Correlation Between Recommendation System Performance (NDCG@5) and Model Training Computational Time for Various News Component Combinations

5. Conclusions

The study was conducted using the Neural News Recommendation with Attentive Multi-View Learning (NAML) model for news recommendation systems. Several methods were explored to generate vector representations of news entities. The method of linking entities to a knowledge base outperformed the baseline by up to 1.53%, 2.89%, 2.23%, and 1.96% in AUC, MRR, NDCG@5, and NDCG@10, respectively. Further experiments demonstrated that the NAML model incorporating named entities led to even greater performance improvements—up to 4.63%, 7.49%, 5.86%, and 5.05% across the same metrics—indicating its effectiveness for news recommendation in the Indonesian language. Moreover, integrating all available news metadata, including tags and named entities, into the NAML model further improved performance, yielding gains of up to 2.06%, 2.93%, 2.39%, and 2.08% for AUC, MRR, NDCG@5, and NDCG@10, respectively.

In conclusion, the NAML method demonstrates strong performance on Indonesian-language news data, making it a promising and effective solution for the news recommendation task. Moreover, the integration of news tags and named entities significantly boosts the performance of the model, indicating the value of leveraging additional semantic information in personalized news recommendation systems. Future research directions include leveraging Large Language Models (LLMs) to generate semantically rich and up-to-date word embeddings, constructing a large-scale and human-validated entity dataset followed by accurate entity linking and embedding using models beyond TransE, exploring various knowledge-aware models for improved integration of entity information, developing fairness-aware models to ensure sentiment diversity in recommendations, and addressing the cold start problem for users without prior interaction history.

Acknowledgement
This research was funded by the Directorate of Research and Development, Universitas Indonesia, under Hibah PUTI Pascasarjana 2024 (Grant No. NKB-28/UN2.RST/HKP.05.00/2024).

References

[1] Davidson, James, Benjamin Liebald, Junning Liu, Palash Nandy, Taylor Van Vleet, Ullas Gargi et al. (2010) The YouTube Video Recommendation System. *Proceedings of the Fourth ACM Conference on Recommender Systems*, Association for Computing Machinery, New York, NY, USA. p. 293–96. https://doi.org/10.1145/1864708.1864770

[2] Das, Abhinandan S, Mayur Datar, Ashutosh Garg and Shyam Rajaram. (2007) Google news personalization: scalable online collaborative filtering. *WWW*, p. 271–80.

[3] Freyne, Jill, Michal Jacovi, Ido Guy and Werner Geyer. (2009) Increasing Engagement through Early Recommender Intervention. *Proceedings of the Third ACM Conference on Recommender Systems*, Association for Computing Machinery, New York, NY, USA. p. 85–92. https://doi.org/10.1145/1639714.1639730

[4] Baskoro, Jatmiko Budi and Evi Yulianti. (2022) SGCF: Inductive Movie Recommendation System with Strongly Connected Neighborhood Sampling. *Jurnal Ilmu Komputer Dan Informasi*, **15**, 55–67. https://doi.org/10.21609/jiki.v15i1.1066

[5] Liu, Yun and Jun Miyazaki. (2023) Knowledge-aware attentional neural network for review-based movie recommendation with explanations. *Neural Computing and Applications*, **35**, 2717–35.

[6] Wang, Hongwei, Fuzheng Zhang, Xing Xie and Minyi Guo. (2018) DKN: Deep Knowledge-Aware Network for News Recommendation. *Proceedings of the 2018 World Wide Web Conference*, International World Wide Web Conferences Steering Committee, Republic and Canton of Geneva, CHE. p. 1835–44. https://doi.org/10.1145/3178876.3186175

[7] An, Mingxiao, Fangzhao Wu, Chuhan Wu, Kun Zhang, Zheng Liu and Xing Xie. (2019) Neural News Recommendation with Long- and Short-term User Representations. *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*, Association for Computational Linguistics, Florence, Italy. p. 336–45. https://doi.org/10.18653/v1/P19-1033

[8] Qi, Tao, Fangzhao Wu, Chuhan Wu and Yongfeng Huang. (2022) News Recommendation with Candidate-Aware User Modeling. *Proceedings of the 45th International ACM SIGIR Conference on Research and Development in Information Retrieval*, Association for Computing Machinery, New York, NY, USA. p. 1917–21. https://doi.org/10.1145/3477495.3531778

[9] Wu, Chuhan, Fangzhao Wu, Suyu Ge, Tao Qi, Yongfeng Huang and Xing Xie. (2019) Neural News Recommendation with Multi-Head Self-Attention. *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP)*, Association for Computational Linguistics, Hong Kong, China. p. 6389–94. https://doi.org/10.18653/v1/D19-1671

[10] Wu, Chuhan, Fangzhao Wu, Mingxiao An, Jianqiang Huang, Yongfeng Huang and Xing Xie. (2019) NPA: Neural news recommendation with personalized attention. *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, p. 2576–84.

[11] Wu, Chuhan, Fangzhao Wu, Mingxiao An, Jianqiang Huang, Yongfeng Huang and Xing Xie. (2019) Neural News Recommendation with Attentive Multi-View Learning. *Proceedings of the 28th International Joint Conference on Artificial Intelligence*, AAAI Press. p. 3863–69.

[12] Wu, Chuhan, Fangzhao Wu, Mingxiao An, Yongfeng Huang and Xing Xie. (2019) Neural News Recommendation with Topic-Aware News Representation. *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*, Association for Computational Linguistics, Florence, Italy. p. 1154–59. https://doi.org/10.18653/v1/P19-1110

[13] Assuncao, Willian G, Lara S G Piccolo and Luciana A M Zaina. (2022) Considering emotions and contextual factors in music recommendation: a systematic literature review. *Multimedia Tools and Applications*, **81**, 8367–8407.

[14] Martijn, Millecamp, Cristina Conati and Katrien Verbert. (2022) ``Knowing me, knowing you'': personalized explanations for a music recommender system. *User Modeling and User-Adapted Interaction*, **32**, 215–52.

[15] Herlocker, Jonathan L, Joseph A Konstan, Loren G Terveen and John T Riedl. (2004) Evaluating Collaborative Filtering Recommender Systems. *ACM Trans Inf Syst*, Association for Computing Machinery, New York, NY, USA. **22**, 5–53. https://doi.org/10.1145/963770.963772

[16] Phelan, Owen, Kevin McCarthy, Mike Bennett and Barry Smyth. (2011) Terms of a feather: Content based news recommendation and discovery using twitter. *ECIR*, p. 448–59.

[17] Okura, Shumpei, Yukihiro Tagami, Shingo Ono and Akira Tajima. (2017) Embedding-based news recommendation for millions of users. *KDD*, p. 1933–42.

[18] Wu, Chuhan, Fangzhao Wu, Yongfeng Huang and Xing Xie. (2023) Personalized News Recommendation: Methods and Challenges. *ACM Trans Inf Syst*, Association for Computing Machinery, New York, NY, USA. **41**. https://doi.org/10.1145/3530257

[19] Kim, Yoon. (2014) Convolutional Neural Networks for Sentence Classification. *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*,.

[20] Lian, Jianxun, Fuzheng Zhang, Xing Xie and Guangzhong Sun. (2018) Towards Better Representation Learning for Personalized News Recommendation: a Multi-Channel Deep Fusion Approach. *IJCAI*, p. 3805–11.

[21] Wu, Fangzhao, Ying Qiao, Jiun-Hung Chen, Chuhan Wu, Tao Qi, Jianxun Lian et al. (2020) MIND: A Large-scale Dataset for News Recommendation. *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, Association for Computational Linguistics, Online. p. 3597–3606. https://doi.org/10.18653/v1/2020.acl-main.331

[22] Kahfi, M.S., E. Yulianti and A.F. Wicaksono. (2023) Incorporating News Tags into Neural News Recommendation in Indonesian Language. *International Journal of Advanced Computer Science and Applications*, **14**. https://doi.org/10.14569/IJACSA.2023.01411124

[23] Meng, Xiangfu, Hongjin Huo, Xiaoyan Zhang, Wanchun Wang and Jinxia Zhu. (2023) A Survey of Personalized News Recommendation. *Data Science and Engineering*,.

[24] Takeda, Yasuhiro, Ikki Ohmukai, Shin ADACHI and Toshikazu Kato. (2009) NEWS RECOMMENDATION METHOD USING GROUP TAGS AND THESAURUSES. *Transactions of Japan Society of Kansei Engineering*, **8**, 813–18. https://doi.org/10.5057/jjske.8.813

[25] Shen, Yan, Pengqiang Ai, Yingyuan Xiao, Wenguang Zheng and Wenxin Zhu. (2018) A Tag-Based Personalized News Recommendation Method. *2018 14th International Conference on Natural Computation, Fuzzy Systems and Knowledge Discovery (ICNC-FSKD)*, p. 964–70. https://doi.org/10.1109/FSKD.2018.8687219

[26] Qi, Tao, Fangzhao Wu, Chuhan Wu, Peiru Yang, Yang Yu, Xing Xie et al. (2021) HieRec: Hierarchical User Interest Modeling for Personalized News Recommendation. *Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing (Volume 1: Long Papers)*, Association for Computational Linguistics, Stroudsburg, PA, USA. p. 5446–56. https://doi.org/10.18653/v1/2021.acl-long.423

[27] Mikolov, Tomas, Kai Chen, Greg Corrado and Jeffrey Dean. (2013) Efficient Estimation of Word Representations in Vector Space. *ArXiv*,.

[28] Bojanowski, Piotr, Edouard Grave, Armand Joulin and Tomas Mikolov. (2017) Enriching Word Vectors with Subword Information. *Transactions of the Association for Computational Linguistics*, MIT Press, Cambridge, MA. **5**, 135–46. https://doi.org/10.1162/tacl_a_00051

[29] Pennington, Jeffrey, Richard Socher and Christopher Manning. (2014) GloVe: Global Vectors for Word Representation. *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, Association for Computational Linguistics, Doha, Qatar. p. 1532–43. https://doi.org/10.3115/v1/D14-1162

[30] Vaswani, Ashish, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N Gomez et al. (2017) Attention is All you Need. In: Guyon I, Luxburg U Von, Bengio S, Wallach H, Fergus R, Vishwanathan S, et al., editors. *Advances in Neural Information Processing Systems*, Curran Associates, Inc.

[31] Gambhir, Mahak and Vishal Gupta. (2022) Deep learning-based extractive text summarization with word-level attention mechanism. *Multimedia Tools and Applications*, **81**, 20829–52.

[32] Yulianti, Evi, Nicholas Pangestu and Meganingrum Arista Jiwanggi. (2023) Enhanced TextRank using weighted word embedding for text summarization. *International Journal of Electrical and Computer Engineering*, Institute of Advanced Engineering and Science (IAES). **13**, 5472–82. https://doi.org/10.11591/ijece.v13i5.pp5472-5482

[33] Rampisela, Theresia V and Evi Yulianti. (2020) Academic Expert Finding in Indonesia using Word Embedding and Document Embedding: A Case Study of Fasilkom UI. *2020 8th International Conference on Information and Communication Technology (ICoICT)*, p. 1–6. https://doi.org/10.1109/ICoICT49345.2020.9166249

[34] Lima, Rennan C and Rodrygo L T Santos. (2022) On Extractive Summarization for Profile-Centric Neural Expert Search in Academia. *Proceedings of the 45th International ACM SIGIR Conference on Research and Development in Information Retrieval*, Association for Computing Machinery, New York, NY, USA. p. 2331–35. https://doi.org/10.1145/3477495.3531713

[35] Nissa, Nuzulul Khairu and Evi Yulianti. (2023) Multi-label text classification of Indonesian customer reviews using bidirectional encoder representations from transformers language model. *International Journal of Electrical and Computer Engineering*, Institute of Advanced Engineering and Science (IAES). **13**, 5641–52. https://doi.org/10.11591/ijece.v13i5.pp5641-5652

[36] Aoumeur, Nour Elhouda, Zhiyong Li and Eissa M Alshari. (2023) Improving the Polarity of Text through word2vec Embedding for Primary Classical Arabic Sentiment Analysis. *Neural Processing Letters*, **55**, 2249–64.

[37] Bordes, Antoine, Nicolas Usunier, Alberto Garcia-Durán, Jason Weston and Oksana Yakhnenko. (2013) Translating embeddings for modeling multi-relational data. *Proceedings of the 26th International Conference on Neural Information Processing Systems - Volume 2*, Curran Associates Inc., Red Hook, NY, USA. p. 2787–95.

[38] Iana, Andreea, Goran Glavas and Heiko Paulheim. (2023) Simplifying Content-Based Neural News Recommendation: On User Modeling and Training Objectives. *Proceedings of the 46th International ACM SIGIR Conference on Research and Development in Information Retrieval*, Association for Computing Machinery, New York, NY, USA. p. 2384–88. https://doi.org/10.1145/3539618.3592062

[39] Desyaputri, Diandra Mayang, Alva Erwin, Maulahikmah Galinium and Didi Nugrahadi. (2013) News recommendation in Indonesian language based on user click behavior. *2013 International Conference on Information Technology and Electrical Engineering (ICITEE)*, IEEE. p. 164–69. https://doi.org/10.1109/ICITEED.2013.6676232

[40] Putri, Wahyuningdiah T. H., Muhammad Singgih Prastio, Retno Hendrowati, Yustiana Sari and Harry T. Y. Achsan. (2019) Content-based Filtering Model for Recommendation of Indonesian Legal Article Study Case of Klinik Hukumonline. *2019 International Workshop on Big Data and Information Security (IWBIS)*, IEEE. p. 9–14. https://doi.org/10.1109/IWBIS.2019.8935726

[41] Wu, Chuhan, Fangzhao Wu and Yongfeng Huang. (2022) Rethinking InfoNCE: How Many Negative Samples Do You Need? In: Raedt L De, editor. *Proceedings of the Thirty-First International Joint Conference on Artificial Intelligence, IJCAI-22*, International Joint Conferences on Artificial Intelligence Organization. p. 2509–15. https://doi.org/10.24963/ijcai.2022/348